Assistive Technology for Visually Impaired and Blind People

Marion A. Hersh
Michael A. Johnson
Editors

with David Keating, Stuart Parks, Gunnar Jansson, Brian S. Hoyle, Dean A. Waters, Nicholas A. Bradley, Mark D. Dunlop, Michael May, Charles LaPierre, Elizabeth M. Ball, Tai Fook Lim Jerry, Han Leong Goh, Kok Kiong Tan, Gareth Evans, Paul Blenkhorn, Rüdiger Hoffmann, James R. Fruchterman, David Crombie and Roger Lenoir

CD-ROM INCLUDED

Springer

Assistive Technology for Visually Impaired and Blind People

Related titles

Assistive Technology for the Hearing-impaired, Deaf and Deafblind
Marion A. Hersh and Michael A. Johnson (Eds.)

Marion A. Hersh and Michael A. Johnson (Eds.)

with David Keating, Stuart Parks, Gunnar Jansson, Brian S. Hoyle, Dean A. Waters, Nicholas A. Bradley, Mark D. Dunlop, Michael May, Charles LaPierre, Elizabeth M. Ball, Tai Fook Lim Jerry, Han Leong Goh, Kok Kiong Tan, Gareth Evans, Paul Blenkhorn, Rüdiger Hoffmann, James R. Fruchterman, David Crombie and Roger Lenoir

Assistive Technology for Visually Impaired and Blind People

Springer

Marion A. Hersh, MSc, PhD
Department of Electronics
and Electrical Engineering
University of Glasgow
Glasgow
G12 8LT
UK

Michael A. Johnson, MSc, PhD
Industrial Control Centre
University of Strathclyde
Graham Hills Building
50 George Street
Glasgow
G1 1QE
UK

ISBN 978-1-84628-866-1

e-ISBN 978-1-84628-867-8

DOI 10.1007/978-1-84628-867-8

British Library Cataloguing in Publication Data
Assistive technology for vision-impaired and blind people
1. People with visual disabilities – Orientation and mobility – Technological innovations 2. People with visual disabilities – Means of communication – Technological innovations 3. Blind, Apparatus for the – Technological innovations 4. Communication devices for people with disabilities – Technological innovations
I. Hersh, Marion A., 1956– II. Johnson, Michael A., 1948– III. Keating, David
362.4’183
ISBN-13: 9781846288661

Library of Congress Control Number: 2007931821

Braille Lite, JAWS®, PAC Mate™, SARA™, StreetTalk™ are trademarks and registered trademarks owned by Freedom Scientific, 11800 31st Court North, St. Petersburg, FL 33716-1805, USA

Windows® is a registered trademark of the Microsoft Corporation, One Microsoft Way, Redmond, WA 98052-6399 USA

SONUS-1XT, XT-1, iVOX, PURE, PURE Digital are trademarks owned by Imagination Technologies, Home Park Estate, Kings Langley, Herts. WD4 8DH, UK

Ultracane™ is a trademark owned by Sound Foresight Ltd., Douris House, Roecliffe Business Centre, Roecliffe, YO51 9NE, UK

Colbolt Speechmaster™ is a trademark owned by Cobolt Systems Ltd, The Old Mill House, Mill Road, Reedham, Norwich, Norfolk, NR13 3TL, UK. The trade mark is applied to a range of products including DialTalk, Sensa Cup Level Indicator, Talking Ear Thermometer

TalkingSigns® is a registered trademark owned by Talking Signs Inc., 812 North Boulevard – Baton Rouge, LA 70802, USA

TalkingLights® is a registered trademark owned by Talking Lights LLC, 28 Constitution Road, Boston, MA 02129, USA

TeleTalk™, DigiCite™, Responder™, Vindicator™, WorkCam™, or SpeakFAST™ are trademarks owned by Compusult Ltd., 40 Bannister Street, Mount Pearl, Newfoundland, Canada, A1N 1W1

LaserCane™ is a trademark owned by Nurion-Raycal Industries Inc., Station Square, Bldg ##2, Suite B, Paoli, Pennsylvania 19301, United States

LeoBraille™ Display, C2 Talking Compass™, Colombus Talking Compass™ are trademarks owned by Robotron Group, 15 Stamford Road, Oakleigh, Victoria 3166, Australia

EZ® and EZ Access® are registered trademarks of the University of Wisconsin-Madison, Madison, WI 53706, USA

BRAILLEX® is a registered trademark of F.H. Papenheimer GmbH and Co., KG, POBox 1620, 58211 Schwerte, Germany

LabVIEW™ is a trademark of National Instruments Corporation, 11500 N. Mopac Expressway, Austin, Texas 78759-3504, USA, http://www.ni.com/

MATLAB® is a registered trademark of The MathWorks, Inc., 3 Apple Hill Drive, Natick, MA 01760-2098, USA, http://www.mathworks.com

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Preface

Until quite recently, the medical model of disability was dominant and assistive technology was viewed as an extension of rehabilitation engineering. But times and viewpoints change so that now social inclusiveness is the pervading ethos of disability legislation, regulations and guidelines. While the existence of new legislative frameworks does not always mean that effective implementation has occurred in the community, it is a beginning. Thus, it is the widespread acceptance of the social model of disability that is driving these changes and it is the tools of assistive technology that are the physical enablers of social inclusiveness.

While we have previously published on Assistive Technology for Hearing Impaired, Deaf and Deafblind People (Springer-Verlag London 2003, ISBN 978-1-85233-382-9), this companion volume strikes out in a new direction by using the social model of disability as a framework. In Chapter 1, we present a comprehensive assistive technology (CAT) model that is designed to provide a generic and holistic description of all aspects of assistive technology whether social, human performance, or the engineering technology used. The idea is that the model can be used to provide the systematic vocabulary and interpretation needed to describe any branch of assistive technology applications. The book itself is structured around the activities module of the CAT model and there are several overview or survey chapters that make recourse to various aspects of the CAT model. Of course the volume concentrates on assistive technology for visually impaired and blind people and the various contributing authors have written about their specific assistive technological contributions to this field.

The objective of systematically reporting on assistive technology for visiually impaired people and also trying to imbue the survey chapters with a descriptive paradigm based on the social model of disability was an ambitious one. We could not have accomplished such a task without the cooperation, enthusiasm and, above all, the patience of our collaborating authors. We should like to thank them all for their help in seeing this publishing project come to fruition. We have been very fortunate to meet some of our collaborators at the Workshops and Conferences on Assistive Technologies for Vision and Hearing Impairment that we organise with invaluable European Union support. In many cases this has given us the opportunity to discuss and debate the engineering issues described in this book.

Patience, too, is a virtue our Springer Engineering editorial staff: Oliver Jackson and Anthony Doyle have in abundance. We should like to acknowledge their enthusiasm and support through the long gestation of this publishing project. Our copy editor, John Kirby, is also to be thanked for producing an elegantly presented volume. Thanks are also due for administrative and graphical support given by Vi Romanes and Peter McKenna of the Department of Electronics and Electrical Engineering at the University of Glasgow during the years of preparation for this volume.

We hope this book with its new modelling perspectives and its systematic coverage of assistive technology will inspire many new projects, new courses, and new ways to secure social inclusiveness for the visually impaired and blind community.

Marion A. Hersh and Michael A. Johnson
Glasgow, Scotland, U.K.

Who should read this book?

This book is designed to inform a wide range of current and future professionals about the assistive technology used by visually impaired and blind people to achieve independence and social inclusiveness in the home and the wider community. Basic engineering principles are explained and the ways these are used to develop and drive assistive technology applications for visually impaired and blind people described. The volume has some chapters that refer to a generic comprehensive assistive technology model to capture the essentials of the applied system and this model should find applications in other assistive technology areas.

The book is suitable for electrical engineering, mechanical engineering and scientific professionals. It is also considered highly appropriate for undergraduate courses in the discipline of assistive technology. Thus, we hope this book will be well placed to meet this need as a course textbook or to supplement existing course material. The authors have been encouraged to see many engineering undergraduates enjoy this type of material and it is hoped that this enjoyment will fire the ingenuity of new generations of engineering students to find new and innovative ways to develop assistive technology for visually impaired and blind people.

An Overview of the Book

The book has a map, for the first four chapters are devoted to fundamentals: disability and assistive technology models, eye physiology and sight, sight measurement principles and technology and finally, haptics. Subsequently groups of chapters explore the topics of mobility, communications and access to information, daily living, education and employment, and recreational activities. These chapter groupings follow the structure of the *Activities* module of the comprehensive assistive technology model as presented in Chapter 1 of the book.

The book is designed so that each chapter is self-contained and can be read on its own, although the overview chapters (Chapters 5, 10, 12, 17 and 18) assume some familiarity with the CAT model material in Chapter 1. Each chapter is motivated by specific learning objectives and contains introductory material or descriptions of the basic principles underlying the technology or applications area. The chapters close with a chapter summary, questions and suggestions for more investigative projects. Full citation details for references to journals, books, and conference papers are given along with information about useful related websites.

A brief description of the contents of each chapter along with full details of the chapter authors can be found next. For the interested reader, biographical sketches of all the contributing authors can be found at the end of the book. These are given in alphabetical order of the author family names. The concept of the book and the overall editorial direction was solely the responsibility of Marion Hersh and Michael Johnson. However, as is usual for a contributed book, the chapter authors are responsible for the opinions and factual accuracy expressed in their particular contributions.

Chapters on Fundamentals

1 Disability and Assistive Technology Systems

Marion Hersh[1] and Michael Johnson[2]

[1]University of Glasgow, [2]University of Strathclyde, Glasgow, U.K.

The social model of disability is discussed highlighting the recent pre-eminence achieved over the medical model of disability. The concept of "quality of life" indices is explored and its relevance to assessing assistive technology applications is described. A survey of the main assistive technology quality of life procedures is presented and the value of the individual procedures considered.

The later sections of the chapter investigate whether assistive technology can be described in a single holistic and generic model, the idea being that the model will provide a uniform and consistent framework for analysing existing applications and for creating or synthesizing new assistive technology systems. The new comprehensive assistive technology (CAT) model is presented and its use demonstrated in these sections.

2 Perception, the Eye and Assistive Technology Issues

Marion Hersh

University of Glasgow, Glasgow, U.K.

Many assistive technology systems for the visually impaired are supported by contributions from the senses of touch, hearing and smell. This chapter opens with a description of the nature of multisensory perception as this forms an important context for the design and use of assistive technology systems.

The chapter then concentrates on the sense of vision. Basic eye physiology is presented along with descriptions of some of the capabilities of the human eye for binocular vision, colour vision and motion tracking.

A brief review of the demographics of vision impairment is given and this is followed by descriptions of the effects of typical vision impairments. A set of photographic images illustrates the conditions described. The basics of simple spectacle provision close the chapter.

3 Sight Measurement

David Keating and Stuart Parks

Gartnavel Hospital, Glasgow, U.K.

Measurement science for the sense of sight has exploited advanced computer technology to emerge as an exciting technical and medical discipline. The chapter presents a full survey of sight measurement methods describing procedures, engineering principles, technological construction and diagnostic motivation. The chapter opens with the classical measurement tests for visual acuity, field of vision, and intraocular pressure, followed by the techniques used in biometry and ocular examinations.

The more advanced technological fields of optical coherence tomography and ocular electrophysiology are described in the last two sections of the chapter. These techniques have developed in sophistication over the last twenty years or so. Advances in computer visualisation software, laser technology, data collection, signal processing algorithms and human-sensor interface systems have all been used to provide complex and accurate measurements and visualisations of the eye physiology and functions for clinical diagnosis. The chapter presents a state-of-the-art review of these sight measurement advances.

4 Haptics as a Substitute for Vision

Gunnar Jansson

Department of Psychology, Uppsala University, Sweden

Historically the sense of touch has been used extensively to generate information for the visually impaired person. This chapter surveys the underlying principles of haptics and the perceptual capabilities of touch achievable with the human hand. After a presentation of these haptic fundamentals, the chapter proceeds to investigate how haptics can be used and enhanced through training or with the aid of specialist tools. A central section of the chapter concentrates on *low-tech* haptic applications; some, like the long cane and the guide dog, are for mobility whilst others, like Braille and embossed pictures, are for information from text, as well as embossed graphics. Subsequent sections in the chapter examine the more technologically advanced applications of haptic science. Of particular importance are the technologies for haptic computer interfaces and for haptic displays. A project to provide haptic access to museum pieces for visually impaired people is one outcome of this advanced work.

Chapters on Mobility and Navigation

5 Mobility: An Overview

Marion Hersh[1] and Michael Johnson[2]

[1]University of Glasgow, [2]University of Strathclyde, Glasgow, U.K.

Visually impaired people usually require assistive technology to aid mobility and retain independent travel within the community environment. This overview chapter opens with a discussion of the travel activity and investigates how people negotiate a desired route or journey. Assistive technology for visually impaired person's travel has had a long history and this is briefly reviewed. One finding is that there have been quite a few attempts to harness the available contemporary technological advances in mobility assistive devices. The subsequent development of the chapter pursues three main topics: obstacle avoidance, navigation and orientation and the design of accessible environments. The presentation reveals that most effort has been devoted to obstacle avoidance assistive technology and that more recently global positioning system and mobile telephone technology has begun to impact the development of viable navigation and orientation assistive technology. The final section of the chapter reviews progress towards the accessible environment that is just beginning to appear in many cityscapes.

6 Mobility AT: The Batcane (UltraCane)

Brian Hoyle and Dean Waters

University of Leeds, Leeds, U.K.

The use of the long cane by visually impaired people as an obstacle detector is long standing. More recently the basic cane design has been equipped with laser or ultrasound transmitters and sensors and an interpretive human interface to improve its effectiveness, the objective being to allow safe travel by a visually impaired person. This chapter reports an important case study of the steps involved in developing an advanced technology obstacle avoidance cane that used bat echolocation signal processing techniques and ultrasonic technology. The final cane design is now marketed worldwide as the UltraCane™.

The chapter reviews the basic technological principles for ultrasonic waves and the advanced signal processing methods used. There is an extended discussion of all the design and construction issues followed by a description of the final engineering and prototype test phase. The chapter closes with an examination of the issues involved in bringing the prototype to eventual commercialisation.

7 Navigation AT: Context-aware Computing

Nicholas Bradley and Mark Dunlop

University of Strathclyde, Glasgow, U.K.

Achieving independence whilst attempting a journey involving long distance navigation still remains a significant challenge for the visually impaired community. In this chapter the contribution that might be made by context-aware computing is explored. The first half of the chapter investigates different aspects of the long distance navigation problem and presents a survey of existing assistive technology, along with an introduction to cognitive maps and navigation learning strategies.

It is in the second half of the chapter that the principles and potential application areas for context-aware computing are introduced. The topics examined include how contextual information, for example, about location and personal preferences, can be embedded into user-computer interactions and how these facilities and capabilities could be used to assist the visually impaired traveller on a long distance journey. The chapter closes with sections on specific prototype applications and some results from the authors' own research tests.

8 Accessible Global Positioning System (GPS) and Related Orientation Technologies

Michael May and Charles LaPierre

Sendero Group, United States of America

One of the better-developed technologies for pinpointing a person's location is the global positioning system (GPS). This US positioning technology has been widely exploited in many consumer applications and over future years alternative systems will become available for use (notably the European Galileo system). The success of a system for use by the visually impaired will depend on the accessibility of the interface design and the value of the information imparted to the user. This case study chapter looks at all the issues from the simple principles of GPS technology through to interface design, development and testing and finally the commercialisation aspects of marketing the end product accessible GPS device and system. The chapter is based on the authors' joint and direct experience of developing and then marketing an accessible GPS product for visually impaired people.

9 Electronic Travel Aids: An Assessment

Elizabeth M. Ball

Ergonomics and Safety Research Institute, Loughborough University, Loughborough, U.K.

Many assistive technology mobility products are expensive and visually impaired people are a community group with considerable variability in their range of sight abilities. Thus, it is very useful to learn about the relative successes and limitations of many of the currently available mobility products; this chapter provides such an assessment.

The chapter opens with an analysis of the various types of methods that can be used to set-up an end-user assessment exercise. Part of this concerns the framework of user requirements of the products to be assessed and another part is concerned with selecting the way of collecting the raw data of end-users responses and experiences.

The second part of the chapter presents the author's findings for an end-user assessment of six obstacle avoidance mobility aids and two accessible navigation aids. The chapter closes with a discussion on the importance of training to achieve the best return from the use of advanced technology to assist in mobility and navigation.

10 Accessible Environments

Marion Hersh[1] and Michael Johnson[2]

[1]University of Glasgow, [2]University of Strathclyde, Glasgow, U.K.

One consequence of the social model of disability is the need for the community environment to be physically accessible to all members of society. This accessibility applies to both the outside environment of, for example, parks, shopping malls and bus stations and to the interior environments of, for example, schools, hospitals, health centres, sports centres, shops, banks, town halls and entertainment complexes. This chapter provides an overview of the types of features that make the community environments accessible for visually impaired people.

The opening section of the chapter looks at the legislative and regulatory frameworks and the general design principles for accessible environments. This is followed by two sections covering the streetscape and buildings respectively. More challenging applications involve embedding detailed information technology modules into the environment and these are covered in the last two sections of the chapter. This is where exemplary applications of accessible public transport and way-finding systems are described.

11 Accessible Bus System: A Bluetooth Application

Tai Fook Lim Jerry, Han Leong Goh and Kok Kiong Tan

National University of Singapore, Republic of Singapore

The flexibility and freedom offered by new wireless technologies are often discussed in the media and on the Internet. But having this potential translated into working assistive technology systems is not so common. In this chapter, an application of Bluetooth technology for a bus alerting and information system suitable for use by visually impaired people is described. The chapter reports case study material across all the activities found in a typical prototype development project.

The chapter opens with a detailed consideration of the elements of Bluetooth technology and has a short comparison section with other competing wireless technologies. Having selected Bluetooth as the enabling technology, the chapter then reports on the design requirements for the bus alerting system. The system development is presented in detail along with careful consideration of the user interface needed for visually impaired people using mobile telephone technology. A discussion of future plans and commercialisation issues closes the chapter.

12 Accessible Information: An Overview

Marion Hersh[1] and Michael Johnson[2]

[1]University of Glasgow, [2]University of Strathclyde, Glasgow, U.K.

In our modern society, increasingly complex media and technology are being used to transmit information. However, to participate and enjoy the benefits of the information revolution requires a continual familiarity with the new developments, so it is important that this area remains accessible to the visually impaired community.

This chapter opens with a review of the principles and technologies of low vision aids that are used to access print. Sections on audio transcription and Braille as access routes to print information then follow. It is the recent developments in speech processing and speech synthesis technology that are drivers in the wider use of audio as an information interface for the visually impaired. Major sections of the chapter describe the accessible computer and the accessible Internet. Both are extremely important in the processing and provision of information and there are many interface options to make these systems accessible to the visually impaired. Finally, since mobile telephony is increasingly accruing computer and Internet capabilities, the chapter closes by reviewing accessible communications technology.

13 Screen Readers and Screen Magnifiers

Gareth Evans and Paul Blenkhorn

University of Manchester, Manchester, U.K.

Two extremely important assistive technologies for the accessible computer are screen magnifiers and screen readers. Together these two tools create the accessibility for computer output needed by a wide range of visually impaired computer users.

The chapter has two major sections, one for screen magnifiers and one for screen readers. Within these extended sections are topics like historical perspectives on the technology developments, the architectures and implementation of the technologies and other sections on particular or special features of these two assistive technology systems.

The chapter is completed by sections on hybrid screen reader-magnifiers, self-magnifying applications and self-voicing applications.

14 Speech, Text and Braille Conversion Technology

Rüdiger Hoffmann

Dresden University of Technology, Dresden, Germany

This chapter is devoted to the fascinating triangle of conversion technologies that arise between text, speech and Braille. These are enabling technologies that allow speech to be converted into text as might happen in the creation of a letter, that allows text to be converted into speech as might happen in the reading of a book for enjoyment and then, the additional steps taking text into Braille for the Braille user. The everyday application of these technologies in making information and computers accessible to the visually impaired is an important assistive technology area.

The chapter opens with a presentation on the fundamentals of speech and text conversion technologies. The spectral analysis of speech is an important theoretical and practical component of this introductory material. Then, this is followed by sections that examine the three technologies; speech-to-text, text-to-speech and Braille conversion in detail. These sections describe the technological principles used and the equipment and applications that follow. The section on Braille conversion has additional material on more specialised Braille applications like Braille refreshable displays, reading machines and access to telecommunication devices.

15 Accessing Books and Documents

James Fruchterman

Benetech, Palo Alto, California, USA

Reading is an essential daily living task, and is crucial for school and work. Whether it is sorting the bills, reading a textbook or the daily newspaper, access to reading is critically important to people with disabilities that prevent easy reading of the printed page. Assistive technology has been created to address these needs and bridge the accessibility challenge to print. One of the first challenges is acquiring the text from the printed page. This need is met through optical character recognition that turns an image of the printed page into an accessible digital text file. In this chapter, the fundamentals of OCR technology and reading machines are described. The new international standard for digital talking books, the DAISY standard, is explored. The critically important move to direct digital access to textbooks and newspapers is projected and a discussion of future technological development closes the chapter.

16 Designing Accessible Music Software for Print Impaired People

David Crombie and Roger Lenoir

DEDICON (formerly FNB Nederland), Amsterdam, The Netherlands

Making music is a pastime enjoyed by many and making printed music accessible to the visually impaired community is a particularly challenging and technically interesting problem. Music notation is a form of symbolic information with its own systematic rules; since it is essentially a visual medium the challenge is to make this system accessible to the visually impaired and blind music player. There are three basic formats: Braille music, talking music (that might use the DAISY standard) and large print music, and the chapter covers all three formats.

The chapter opens with a brief introduction to music notation that defines many key terms and structures for later use. This is followed by a survey of the three music formats designed to support accessibility. Some international project activities are described along with attempts to prescribe an international standard for the field. A key section in the chapter carefully details the steps in the production of Braille music and also gives further information about Talking Music. Since these accessibility topics are still developing, closing sections of the chapter review the likely future developments for the field.

17 Assistive Technology for Daily Living

Marion Hersh[1] and Michael Johnson[2]

[1]University of Glasgow, [2]University of Strathclyde, Glasgow, U.K.

The first of the contextual activity areas in the comprehensive assistive technology model is that of Daily Living, and this is the subject of this chapter. Thus, the chapter describes assistive technology solutions designed to remove barriers to enable visually impaired and blind people enjoy independent living in their own home. These assistive technology solutions range from some very simple low technology devices to very sophisticated and specialised high technology solutions giving the chapter a large number of subsections and coverage of a wide variety of engineering methods.

Within this diversity of assistive technology applications, there are some techniques that span several of the activities found in the daily living category. Labelling is one generic assistive technology described in the chapter and this has a range of applications in personal care, food preparation and using appliances. Similar generic technologies are those for light and colour detection and identification and these too are described in this chapter. Along with these general solution technologies, the chapter presents groups of devices for the areas of personal care, time-keeping, alarms and alerting, food preparation, using appliances and money amongst others.

Summary conclusions, projects to pursue and reference citations close the chapter.

18 Assistive Technology for Education, Employment and Recreation

Marion Hersh[1] and Michael Johnson[2]

[1]University of Glasgow, [2]University of Strathclyde, Glasgow, U.K.

Education and employment, and recreational activities complete the trio of contextual activities of the comprehensive assistive technology model, and this final chapter of the book reports on the assistive technology available in these two areas.

Education and employment is divided into six subsections covering learning and teaching, and then five employment activity areas. The generic assistive technology to support mobility, access to information and communication that are essential to access education and employment have been described earlier in the book and in this chapter access to the higher level activities are discussed. For example, this part of the chapter covers access to mathematics in education, and the use of specialised tools in the skilled and non-skilled trades. There is also a discussion of the general levels of access attained by visually impaired and blind people to education and employment in the introductory section of the chapter.

Recreational activities are essential to personal wellbeing and this is the third contextual activities category in the CAT model. This category also divides into six sub-activity areas, and a wide variety of assistive technology solutions is described. These range from an infrared cinema audio description system, through to accessible football and tactile tape measures to facilitate craft activities.

Summary conclusions, projects for further investigation and reference and resource citations close the chapter.

Contents

1 Disability and Assistive Technology Systems

Learning Objectives

Significant changes have taken place in the last decade in the position of disabled people in our societies and communities. A combination of factors has led to new legislation, regulations and standards in many countries to attempt to remove existing discrimination against disabled people and enable them to participate fully in education, employment and the community. These factors include increasing awareness of the benefits to society as a whole of increasing the involvement and independence of disabled people and increasing assertiveness and activity by organisations of disabled people. Implementation of this legislation will involve the further development of assistive technology, as well as better information about what is currently available. This will require an effective and ongoing dialogue between the disabled end-user community and the various professionals, including engineers, technicians, occupational therapists, social workers and medical practitioners, involved in developing and providing assistive technology. In order to support this dialogue there is a need for common definitions and models describing the framework for assistive technology systems. Developing and presenting this framework, which will allow analysis and synthesis of assistive technology systems, as well as appropriate matching to potential end-users, is the main aim of this chapter. This modelling framework is developed within the social model of disability.

Although the medical model, which is closely associated with rehabilitation engineering for people with physical and cognitive impairments, is still more commonly used, it is the social model with a focus on removing social barriers and discrimination rather than rehabilitation that is preferred by organisations representing disabled people. Disabled people and their organisations have had a significant role in achieving the social and attitudinal changes which have led to recognition of the necessity, as well as the social benefits of full social inclusion of disabled people and accessibility of the social infrastructure. This has resulted in new legislation on accessibility and integration. However, much of this legislation is still strongly influenced by the medical model. For example, the definition of a disabled person used in the UK Disability Discrimination Act 1995 is a medical one. Thus social and medical models, as well as the social role of assistive technology are described in Section 1.1 of this chapter.

It is useful to have some means of measuring the effectiveness of assistive technology in overcoming social and infrastructural barriers and improving the quality of life of disabled people. Although not a complete answer, quality of life assessment, sometimes through the use of an index, has a significant role to play. However, it should be noted that there is an important subjective component, which is more difficult to assess, to (disabled) people's experiences and quality of life. Developments in the area of quality of life assessment are described in Section 1.2 of this chapter.

Section 1.3 presents the background to the assistive technology modelling framework developed by the authors. It includes discussion of a number of existing models, including the human activity assistive technology (HAAT) model due to Cook and Hussey (2002), as well as discussion of existing approaches to modelling human activities. A list of criteria to be met by the modelling framework is also presented.

The comprehensive assistive technology (CAT) model developed by the authors is presented in Section 1.4 and some applications of its use are given in Section 1.5. This modelling framework is based on further developments of Cook and Hussey's human activity assistive technology (HAAT) modelling approach. The CAT model represents a holistic paradigm that can be used for assistive technology analysis and synthesis (development of specifications for new devices), as well as for matching assistive technology to a particular end-user. It also has pedagogical value, since it provides a common descriptive framework to support discussion and dialogue within all the disciplines involved in developing and providing assistive technology. The chapter closes with Conclusions in Section 1.6.

The learning objectives for this chapter are:

- Understanding the genesis and content of the social model of disability.
- Appreciating the nature and uses of quality of life assessment.
- Reviewing the different modelling approaches that have been used for modelling assistive technology systems.
- Understanding the structure and uses of the CAT model for assistive technology systems.

1.1 The Social Context of Disability

Since the intended users of assistive technology are disabled people, the definition of assistive technology depends on the definition or model of disability being used. Therefore, before considering definitions of assistive technology, the different models of disability will be presented. There are two main approaches – the medical and social models.

The medical model is based on the international classification of "impairment", "disability" and "handicap" (sometimes referred to as the ICIDH model) developed by the World Health Organisation (WHO) in 1980 (WHO 1980). The WHO defined "impairment" as "any loss or abnormality of psychological, physical or anatomical structure or function". A "disability" then occurs when the impairment prevents a person from being able to "perform an activity in the manner or within the range considered normal for a human being". Hence a "handicap" results when the person with a disability is unable to fulfil their normal role in society and the community at large. Thus the medical model views disability as residing in the individual and focuses on the person's impairment(s) as the cause of disadvantage leading to the approaches of occupational therapy and rehabilitation. It should be noted that organisations of disabled people dislike the term "handicap" and it should not be used.

The social model of disability emphasizes the physical and social barriers experienced by disabled people (Swain *et al.* 2003) rather than their impairments and considers the problem to be in society rather than the disabled person. It is compatible with the empowerment of disabled people and user-centred and participative design approaches (Damodaran 1996; Rowley 1998). The social model was first developed by the Union of the Physically Impaired Against Segregation (UPIAS 1976) and then modified by the Disabled Peoples International (DPI) (Barnes 1994). The model is based on the two concepts of impairment and disability. "Impairment" is defined as the functional limitation caused by physical, sensory or mental impairments. "Disability" is then defined as the loss or reduction of opportunities to take part in the normal life of the community on an equal level with others due to physical, environmental or social barriers.

To illustrate the difference in focus of the two models, the medical model identifies the disability of a partially sighted person as related to their inability to read standard sized print whereas the social model identifies their disability as a consequence of the fact that, for instance, only some books are available in large print versions. Or, more simply, in the social model it is the steps that are the problem not the wheelchair. Organisations of disabled people influenced by the social model of disability have had an important role in changing attitudes towards disabled people and accessibility in the community and securing new services and rights for disabled people. However, considerable further work remains to be done to secure full rights for disabled people, most professionals are still influenced by the medical model and many services and facilities are still provided within a rehabilitation framework.

The importance of the social model was recognised in an update of the WHO classification system. In the new version, commonly termed ICIHD2, the terms "impairment, disability and handicap" were replaced by "disability, activity and participation" (WHO 2001). This model considers disablement to be the result of the interaction between an individual's health and contextual factors. However, it is still the individual's condition rather than external factors that is the main driver of the classification. This differs from the social model in which impairment is considered simply to be part of human diversity, but disability is recognised as being created by social and community environments that have been designed without taking the needs of disabled people into account.

The social model of disability is generally more appropriate for the research, design and development of assistive technology, as the aim of assistive technology products and devices should be to extend opportunities and break down barriers.

At the same time it is useful to be aware of the physical effects of impairments in a context of diversity and equality. This means recognition of physical and other differences, but that everyone is entitled to equal rights and equal opportunities. In addition it should be recognised that there is not a particular norm or way of being human that is better than others, but that the full range of diversity is equally valid and valuable.

The social model of disability can also be used to identify the following two areas of responsibilities for engineers and designers:

1. Design for all; that is designing and constructing devices and environments to be accessible and usable by as wide a range of the population as possible, including disabled people.
2. Design of assistive technology systems, for example, the design of devices to overcome existing environmental and social barriers, thereby extending the opportunities and options open to disabled people.

It should be noted that design for all is much wider than design to include disabled people and it aims to include all (or as many as possible) of the different groups in the population regardless of factors such as age, gender, size, ethnic origin and disability. This is compatible with the social model of disability which looks at removing barriers, since design for all seeks to remove barriers for a wide range of social groups in addition to disabled people.

1.2 Assistive Technology Outcomes: Quality of Life

In terms of the social model the aim of assistive technology is to overcome the gap between what a disabled person wants to do and what the existing social infrastructure allows them to do. It consists of equipment, devices and systems that can be used to overcome the social, infrastructure and other barriers experienced by disabled people and that prevent their full and equal participation in all aspects of society. Measurement of the impacts and outcomes of a particular technology can be used to determine whether the desired benefits have been achieved and to inform further developments. In the case of assistive technology, improved outcome monitoring could have an impact on improving the device-user match and the quality of service provision, which is particularly important in view of the evidence of high rates of device abandonment (Phillips and Zhau 1993). It could also increase accountability (DeRuyter 1997) and provide additional evidence to support demands for increased public funding for assistive technology.

However, relatively little attention has been paid to the assessment of assistive technology outcomes for a number of reasons, including the belief that its benefits are obvious, a focus on the performance of the technology, rather than the interaction of the user with the technology and relatively little demand from stakeholders (Fuhrer *et al.* 2003). As measurement of assistive technology outcomes becomes more common, it will become increasingly important that the focus is on the production of relevant results rather than completing routine documentation and that the procedures involved are not time consuming for either assistive technology users or practitioners. This section will discuss the existing techniques for measuring the quality of life resulting from assistive technology use, how useful these quality of life measures are, and whether additional measures should be developed.

1.2.1 Some General Issues

Measurement of assistive technology outcomes requires consideration of the combined human *plus* assistive technology system, rather than just the performance of the technology on its own. Outcome measures should be designed for the contexts in which assistive technology is actually used. They should also take account of the facts that individuals frequently make use of several assistive technology devices and that the impacts of these devices depend on the application context and type of assistive technology as well as end user characteristics (Gelderblom and de Witte 2002). Outcomes measures should be adaptable to allow measurement of both the specific impacts of a particular assistive device on the quality of life of a given disabled person and the impacts of a particular technology or type of technology on a given group of disabled people.

Both objective and subjective measurements are relevant. Objective measurements include measures of the performance of human assistive technology systems and services, such as the time taken by the user to read a data set or find a particular webpage using a screen reader. Subjective measures include subjective assessments of the user's satisfaction with the device and services and subjective measures of any resulting change in quality of life. In some cases it may be difficult to isolate changes due to the assistive device from those due to other causes (Smith 1996), such as a change in circumstances. Thus it may, for instance, be difficult to determine how much an improvement in quality of life is due to the provision and regular use of an assistive device and how much is the result of a move to a new house closer to family and friends, an improvement in health or getting a job.

Studies on subjective quality of life are generally based on surveys. The nature of the survey process means that it is highly unlikely that all the respondents will complete all the questions. However, the issue of how responses with missing answers to some questions should be treated have not really been resolved. Frequently such responses are deleted, but this can reduce the sample size significantly, for instance by 18.3% for a data set of 10 variables with 2% of the data missing at random (Kim and Curry 1977). Another important issue is respondent and selection bias. Selection bias can probably be avoided by following an appropriate methodology. However respondent bias could be a problem and further research may be required to determine whether people who are satisfied with their assistive devices are more likely to reply to quality of life questionnaires than those who are dissatisfied or who have abandoned them.

There is also a need to use several sources of information and triangulate methodologies to increase the reliability and validity of assessments. The development of quality of life indicators should also involve disabled people to ensure that indicators cover the areas they consider to be important, which may not be the same as those considered important by researchers (Hughes *et al.* 1995).

1.2.2 Definition and Measurement of Quality of Life

A definitive definition of quality of life has not yet been obtained. For instance, a review of 87 studies from diverse literature found 44 different definitions (Hughes *et al.* 1995). Any measurement system for quality of life should ideally be based on an established body of theory and supported by empirical evidence. However, theory in this area is not very well developed, though it has been suggested that Maslow's concept of a hierarchy of needs could provide a suitable framework (Maslow 1968; Sawicki 2002).

There is a considerable body of work on health related quality of life which is based on a medical model of disability and either tacitly or explicitly assumes that impairment reduces quality of life. However, studies show that satisfaction with life increases with the extent of social integration, employment and mobility and is not correlated with the degree of (physical) impairment (Fuhrer *et al.* 1992; Fuhrer 1996). This is in accordance with the social model of disability, which stresses the need to overcome social barriers and social discrimination. However, since the measurement of health related quality of life was developed before specific outcome measures for assistive technology, it will be discussed in Section 1.2.3.

Other approaches, which will not be considered here, have developed from increasing awareness of the importance of sustainable development, as well as recognition of the inadequacy of economic measures such as gross domestic product (GDP) (Hersh 2006). Indicators of this type have been used in public policy making, for instance in measuring areas of deprivation, sustainability and regional economic development in the U.K. (Sawicki 2000). There may therefore be a need for equivalent measures of the impact of assistive technology to support public policy making in this area.

Two of the approaches to defining quality of life for disabled people using assistive technology will be considered here. In the first, Hughes *et al.* (1995) obtained a consensus list of the following 15 dimensions of quality of life from a study of the disability literature:

1. Social relationships and interaction
2. Psychological wellbeing and personal satisfaction
3. Employment
4. Self-determination
5. Autonomy and personal choice
6. Recreation and leisure
7. Personal competence
8. Community adjustment and independent living skills
9. Community integration and normalisation
10. Support services received
11. Individual and social demographic indicators
12. Personal development and fulfilment

13. Social acceptance, social status and ecological fit
14. Physical and material wellbeing
15. Civic responsibility

The second approach due to Schalock (1996) has the following eight dimensions:

1. Emotional wellbeing
2. Interpersonal relations
3. Material wellbeing
4. Personal development
5. Physical wellbeing
6. Self-determination
7. Social inclusion
8. Rights

The second approach has the advantages of being more compact and solely concerned with the impacts on the person, rather than how these impacts are obtained, for instance through work, leisure, civic participation or relationships. It also seems to avoid an explicit or implicit assumption that impairment automatically reduces quality of life, whereas the categories of personal competence and support services received are less likely to be included in a quality of life assessment for non-disabled people. However, the second approach may require an additional category of social identity and status. In addition further research would be required to ensure that the way in which the categories have been framed is culturally neutral and therefore appropriate for use in a wide range of different countries and cultures.

1.2.3 Health Related Quality of Life Measurement

Quality of life is frequently referred to in healthcare literature and quality of life measurements are increasingly used as an endpoint in clinical trials and intervention studies (Bowling 1995; Bowling and Windsor 1999). Health related quality of life studies focus on the measurement of symptoms and function, including levels of impairment, with the aim of assessing health status as part of a clinical measurement of outcome. Unfortunately, this approach with its focus on fulfilling 'normal' roles leads to an automatic devaluing of groups such as disabled or unemployed people and the assumption that disabled people cannot have a good quality of life, which is not in fact the case (Doward and McKenna 2004). Since the introduction of quality of life research in the healthcare context in the mid-1970s, the volume of publications has grown exponentially and there are now over a thousand quality of life citations each year (Fuhrer 2000).

There is still considerable disagreement about the definition and measurement of quality of life (Beckie and Hayduk 1997). However, this has generally been based on traditional health-status measures, under the assumption that they are equivalent to quality of life, and a functional perspective in relation to the ability to perform daily living activities, with disability both considered and measured as something negative (Bowling and Windsor 1999). For instance, a 1994 study of a random sample of 75 articles (Gill and Feinstein 1997) found that most of the assessments used areas of functioning consistent with the biomedical model (Fuhrer 2000), imposing a value system on respondents and deciding independently of them which areas of life were worth being measured. This criticism is supported by evidence of a lack of agreement between individuals' ratings of their own health-related quality of life and those of health care providers or relatives (Slevin *et al.* 1988; Spangers and Aaronson 1992).

An approach, which is more compatible with the social model of disability and empowering to disabled people would replace the idea of unaided functioning by consideration of independence and autonomy. Here independence is understood in the sense of 'control of their life and choosing how that life is led…(and) the amount of control they have over their everyday routine' (Brisenden 1986). Autonomy is defined as the ability to plan one's own life, to enter into relationships with others and together actively participate in the construction of society. These definitions are applicable to both disabled and non-disabled people. A non-disabled person can be non-autonomous if he or she has trouble in planning his or her life, whereas a disabled person who uses assistive technology or a personal assistant will be autonomous if this technology and/or assistant enable him or her to plan his or her life, enter into relationships and participate in society. Similar comments hold for independence.

Studies also indicate little correlation between individuals' 'subjective' judgements of wellbeing and 'objective' measures of income, educational attainment and health status (Diener 1984; Eid and Diener 2004). Health related quality of life measures have the further disadvantage of being formulated in the context of people who are considered 'patients', that is, have been receiving treatment for a health complaint. However, many disabled people are not receiving any medical treatment and, for those who are, this is by no means their whole identity.

Subjective quality of life (sometimes called wellbeing) has been defined as 'the degree to which people have positive appraisals and feelings about their life, considered as a whole' (Fuhrer 2000). Specific measures include the Patient Generated Index (PGI) and the Satisfaction with Life Scale. In the PGI, respondents are first asked to list the most important areas of their life affected by their impairments(s) and then to evaluate the effect in each area and weight the relative importance of the different areas (Fuhrer 2000). Therefore, although the person defines the areas to be considered, the approach is very much based on the medical model of disability.

The quality of life instruments database currently includes 1000 items, with full descriptions given for 454 of them (WWW1 2006). The use of the term instrument indicates a system for the measurement or assessment, in this case of quality of life, which could include a mixture of quantitative and qualitative measures. However, the majority of the instruments are medically based and seem to relate quality of life to wellness, lack of disability and/or functionality. Many of the instruments are specific to populations with particular illnesses or impairments. There seem to be few instruments in this database based on a wider understanding of quality of life or which are appropriate for a general population, including both disabled and non-disabled people and people in different states of health. This is consistent with the tendency of some researchers in the area to use the terms quality of life and general health status interchangeably and to assume that a multiple item health status questionnaire can provide a satisfactory measure of quality of life (McDowell and Newell 1996).

Another approach used in the medical context defines quality of life as the gap between an individual's hopes and expectations and their actual experiences (Calman 1984). This approach has the advantage of including subjective elements and being defined by the individual rather than the researcher. However, there is the associated disadvantage that hopes and expectations are difficult to measure and a preference for the use of definitions which can be quantified and measured. Quality of life scales based on the Calman approach have been developed and include the Schedule for the Evaluation of the Individual Quality of Life (O'Boyle *et al.* 1992, 1993), which is also available in a shorter form, SEIQoL-DW (Browne *et al.* 1994; Hickey *et al.* 1996). In this shorter form, the quality of life scale is uniquely defined for each individual, as it is based on the five quality of life domains that they consider the most important (Mountain *et al.* 2004).

1.2.4 Assistive Technology Quality of Life Procedures

Six different assistive technology specific outcome measures and one ongoing project are reviewed and discussed. These different approaches have been influenced to different extents by health quality of life approaches and a universal standard approach has not yet been developed. Although most of the measures and procedures have been tested both through studies of their performance with different groups of end-users and by experts, there is a need for more comparative studies of the different methods.

1.2.4.1 Life-H

Life-H aims to measure the quality of social participation in terms of the manner in which daily activities and social roles, called life habits, are carried out and with regards to any disruption of these life habits. It uses measures of social participation based on the concept of 'life habits', defined as 'habits that ensure the survival and development of a person in society throughout life' (Noreau *et al.* 2002). It includes essential activities, such as eating and sleeping, activities which are carried out daily, such as personal hygiene and getting out of bed, and other activities based on a mixture of choices and available options, such as social activities, interpersonal relationships and employment. These life habits have been divided into twelve categories (Fouygerollas *et al.* 1998).

The LIFE-H assessment consists of two different questionnaires, a short version for general screening and a longer version for more detailed assessment of specific areas of social participation. The assessment is based on the level of difficulty in performing a life habit (activity) and the type of assistance, for instance an assistive device or human assistance, required and marked on a scale with a continuum of 10 levels. Levels of satisfaction with the accomplishment of each activity are measured on a five-point scale. Activities which are not part of a person's lifestyle from choice are excluded from the assessment. The questionnaire can either be self-administered or form part of an interview. The time required is 30–60 min for the short form and 20–120 min for the long form, depending on the number of categories of life habits being investigated. There is also a version for children (5–13 years old) where life habits considered irrelevant to children, such as parental roles, sexual relationships and employment, have been eliminated.

LIFE-H can be used to investigate the impact of assistive technology by comparison of assessments with and without the technology. However, its use as an outcome measure has been relatively limited. It is available in French and English versions, with plans for translation into Spanish, German, Dutch and Italian. Modifications may be required to ensure that it is appropriate in all countries worldwide (Fouygerollas *et al.* 1998).

1.2.4.2 OT FACT

OT (occupational therapy) FACT (Smith 2002) is a software-based assessment approach to measuring 'function' rather than quality of life. It was originally intended to be used by occupational therapists and aimed at measures of 'functionality' in terms of both observed 'performance' and subjective scoring by the user of their satisfaction with this performance. Question branching is used to investigate areas in which the respondent experiences difficulties or barriers (unfortunately referred to in the literature as 'deficits'). Only issues scored '1' to indicate some barriers are investigated further. The other two possibilities are '0' for no barriers and '2' for impossible to overcome barriers. A version called time series concurrent differential (TSCD) methodology is used to investigate the impact of assistive technology by comparing performance with and without assistive technology. The question branching approach is useful. However, both the software and the inherent philosophy based on 'deficits' are becoming dated.

1.2.4.3 Psychosocial Impacts of Assistive Devices Scale (PIADS)

The PIADS (Day and Jutai 1996; Day *et al.* 2002; Jutai and Day 2002) is a 26-item self-report questionnaire that aims to assess the effects of assistive device use on functional independence, wellbeing and quality of life. It is divided into three sub-scales:

1. Competence (12 items), which measures feelings of competence and efficacy and includes questions on competence, productivity, usefulness, performance and independence.
2. Adaptability (6 items), which measures willingness to try out new things and take risks, and includes questions on participation, willingness to take chances,

eagerness to try new things and the ability to take advantage of opportunities. It is intended to be sensitive to the enabling and liberating aspects of assistive technology.

3. Self-esteem (8 items), which measures feelings of emotional health and happiness and includes questions on self-esteem, security, sense of power, and control and self-confidence. It is intended to be sensitive to the perceived impact of assistive technology on self-confidence and emotional wellbeing.

Responses to each question can range from −3 (most negative impact) to +3 (most positive impact), with zero representing no perceived impact. The questionnaire can generally be completed in 5–10 min. PIADS is currently available in English and translation into other languages, as well as production of versions for children and people with cognitive impairments is being investigated.

1.2.4.4 Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST 2.0)

QUEST 2.0 is intended to evaluate satisfaction with a wide range of assistive technology (Demers *et al.* 2002a,b), expressed in a linear general framework (Simon and Patrick 1997). Satisfaction is defined in QUEST as 'a person's critical evaluation of several aspects of a device' and may be influenced by expectations, perceptions, attitudes and personal values. Expressed satisfaction is considered a reaction to assistive technology provision and may also affect behaviour, for instance through the use of or abandonment of an assistive device.

The initial version of QUEST consisted of 24 items, with responses on 5-point importance and satisfaction scales from 'of no importance' to 'very important' and 'not satisfied at all' to 'very satisfied' respectively. As a result of field-testing (Demers *et al.* 1999a,b), the importance scale was removed, as it did not discriminate reliably between and among assistive technology users and the questionnaire was reduced to 12 items rated on a 5-point satisfaction scale. The items are divided into the following two scales (Demers *et al.* 2000):

- Device scale, consisting of the following eight items: comfort, dimensions, simplicity of use, effectiveness, durability, adjustments, safety and weight.
- Services scale, consisting of the following four items: professional service, follow-up services, repairs/servicing and service delivery.

The scores over all the items on each of these scales, as well as over all the items, are averaged to produce average scores for device satisfaction, satisfaction with the associated services and total satisfaction. QUEST was originally developed in Canadian English and French and has been translated into Dutch, Swedish, Norwegian, Danish and Japanese.

1.2.4.5 Matching Person and Technology (MPT)

MPT is an assessment procedure for use in determining outcomes and the appropriate assistive technology for a particular person in a given environment (Scherer and Craddock 2002). It is based on the matching person and technology model (Scherer 2000; Scherer and Cushman 2000), which is divided into the three main components of the person using the technology, the technology and the milieu or environment.

Two versions have been developed in Ireland and the U.S., with the U.S. version translated into French and Italian. There is also a version for children under 15, known as the matching assistive technology and child (MATCH) version. The procedure is generally performed with the technology user and provider working together. A range of different types of assessments is available and personnel who are trained and experienced in the procedure can carry out assessments in 15–45 min.

The following six-step procedure is used, with the first three steps relating to the questionnaires and the last three to discussion of outcomes and the resulting action to be taken:

1. Using the MPT worksheet to determine 'limitations' in areas such as communication and mobility and initial goals of the service provider and end-user. Potential interventions and the technologies required to support the goals are indicated on the form.
2. The Technology Utilisation Worksheet is used by the end-user and service provider together to record technologies used in the past, satisfaction with them and the technologies the user wants or needs, but has not yet received.
3. The end-user completes their version of the appropriate form for the type of technology (general, assistive, educational, workplace or healthcare) being considered. Alternatively, if appropriate, an oral interview is carried out. The provider completes their version of the same form and identifies any differences between the two forms. The assistive technology device predisposition assessment (ATDPA) asks users about their subjective satisfaction in functional areas (9 items), asks them to prioritize the aspects of their lives they most want to improve (12 items), profiles their psychosocial characteristics (33 items) and asks for their views on 12 aspects of using a particular type of assistive device. The service provider's form aims to enable the service provider to evaluate incentives and disincentives to the user using a given device.
4. The service provider discusses factors with the user that could indicate problems with acceptance or appropriate use of the technology.
5. The user and service provider discuss specific intervention strategies and draw up an action plan to address the problems.
6. The strategies and action plan are written down to increase their likelihood of being implemented and, if required, to provide documentary support for requests for funding or release time for training.

All the forms, except the healthcare form, have two versions for the service provider and end-user respectively, which are intended to be used together to identify characteristics of the person, environment or technology that could lead to inappropriate use or abandonment. The user is required to focus on current feelings and attitudes while completing each form. The user form can be used as the basis for an interview if this is considered more appropriate.

1.2.4.6 Individually Prioritised Problem Assessment (IPPA)

IPPA (Wessels *et al.* 2002) is intended to assess the effectiveness of the assistive technology provision by determining the extent to which the problems and barriers encountered by the user in daily activities have been reduced. The assessment is based on asking the respondent to identify up to seven barriers in carrying out everyday activities that could possibly be countered through the use of assistive technology. This should take place early in the service delivery process to avoid the user being influenced by service providers. A checklist of daily activities can be used to help identify barriers. For each issue an IPPA form is completed and the respondent identifies the importance of the activity and the level of difficulty in carrying it out using a multiple choice score with five values from 'not important at all' to 'most important' and 'no difficulty at all' to 'too much difficulty to perform the activity'. An averaged weighted 'difficulty' score is then obtained, giving a measure of the extent to which the respondent experiences barriers in carrying out daily activities. The process is repeated a few months after the provision of assistive technology to obtain a new 'difficulty' score for the previously identified activities and weights. The original version was produced in English and has been translated into Italian, Dutch, Norwegian and Swedish.

1.2.4.7 Efficiency of Assistive Technology and Services (EATS)

EATS is an ongoing project with the following three main objectives:

- To develop a methodology based on a holistic approach, including end-user preferences, quality of life and social values, for a comprehensive assessment of assistive technologies and services.
- To support decision making about assistive technologies and services through the dissemination of project results.
- To compile and develop a methodology for the comprehensive assessment of assistive technologies and services, including user benefits, quality of life and costs.

1.2.5 Summary and Conclusions

Measurement of assistive technology outcomes could be used to improve the process of matching individuals to particular technologies, as well as to inform policy decision making and demands for increased public funding for assistive technology. There is therefore a need for measurement procedures that are widely applicable and relatively simple to apply. It is also important that the process does not degenerate into a bureaucratic form filling exercise.

Assistive technology outcomes have both objective and subjective components, with the subjective component particularly important. Since one of the potential benefits of assistive technology is improvement of quality of life, outcome measurements frequently focus on quality of life. The measurement of quality of life in the medical context considerably predates the development of specific outcome measures for assistive technology. However, it has the disadvantage of a narrow focus on health and wellness rather than a wider understanding of quality of life, and using measures of unaided functioning that automatically classify disabled people as having a reduced quality of life due to their impairments.

Six different assistive technology specific outcome measures, as well as an ongoing project, have been discussed. All the approaches have advantages and disadvantages and a universal standard has not yet been developed. The different approaches have been influenced to different extents by health related quality of life approaches and associated assumptions that disabled people have a reduced quality of life. However, in most cases such inappropriate assumptions could be edited out by changes in the wording. There may also be advantages in several assessment procedures with different areas of application continuing to be used rather than a *de facto* standard being developed. Most of the measures have been tested both through studies of their performance with different groups of end-users and by experts. However, there is a need for more comparative studies of the different measures, as well as studies outside Europe and the U.S.A.

The current assessment procedures are available in English and sometimes also other European languages and have been developed in Europe or the U.S.A. There is therefore a need for further research to modify these instruments or develop others for use in other parts of the world or even worldwide, as well as to translate the instruments into a much wider range of different languages. It would be useful to have a global standard instrument that is universally applicable and could be translated into any language. However, this may not be feasible, due to differences in culture, lifestyle and attitudes in different parts of the world. Disabled end-users of assistive technology should have a much more central role in future research on outcome measures to ensure that the approaches are relevant to them as well as to the professionals involved in providing assistive technology.

Assistive technology has a potentially important role to play in increasing quality of life by opening up new opportunities to disabled people and increasing the range of options open to them. This, then, is the subject of this book, the development of assistive technology to be used by blind and visually impaired people to increase their options and opportunities.

1.3 Modelling Assistive Technology Systems

There are a number of different definitions of assistive technology. The definition used here will be based on the social model of disability. Other commonly used definitions of assistive technology are based on the medical model of disability.

In terms of the social model the aim of assistive technology is to overcome the gap between what a disabled person wants to do and what the existing social infrastructure allows them to do. It consists of equipment, devices and systems that can be used to overcome the social, infrastructure and other barriers experienced by disabled people that prevent their full and equal participation in all aspects of society. Assistive technology is used in a social, cultural, political, economic and environmental context. This context may facilitate the development and use of assistive technology, pose barriers and constraints or be neutral. Although the infrastructure for the provision of assistive technology is more developed in the industrialised countries, assistive technology is required in the so-called developing countries as well and should be available worldwide. Users and potential users of assistive technology also vary greatly in their characteristics, interests, skills, values and impairments. Assistive technology is also required for a wide range of different types of tasks and applications.

As a result of this great diversity in (potential) end-users, applications and context there is a strong need for a simple, effective and unified modelling framework to support the ongoing dialogue between end-users, therapists, social workers, funding bodies and the engineering community and other researchers and professionals working in the area of assistive technology. A unified modelling framework and terminology will also enable the common structure of assistive technology systems to be understood by different professional domains and provide a pedagogical basis for education and training courses in the assistive technology field. The objectives for the modelling framework can be stated as follows:

1. It should be applicable to any assistive technology system.
2. It should be possible to use the modelling framework to provide a classification of assistive technology systems.
3. The modelling framework should reveal the generic structure of assistive technology systems and lend itself to analysis and synthesis (device specification) procedures.
4. The modelling framework should support the development of new assistive technology systems to meet particular needs.
5. The modelling framework should be able to support the process of providing assistive technology for a particular user, with the aim of avoiding device rejection and abandonment.
6. The modelling framework should give engineers and other professionals a clearer understanding of how assistive technology systems function in a social context.

1.3.1 Modelling Approaches: A Review

Real-world situations are frequently highly complex. This means that simplifying assumptions need to be made to give a model of appropriate size and complexity to analyse, programme or use in predictions. It also frequently means choosing the elements of the situation that are relevant to the particular problem for modelling and ignoring the others. This will generally give a choice of a number of different possible models and modelling structures depending on the aims and desired outcomes of the modelling process. Approaches in the literature to modelling assistive technology have often had different aims or a subset of the aims of the above comprehensive modelling framework. A thorough review of the literature reveals both that the development of assistive technology has only been studied by a limited number of researchers and that there are no existing modelling approaches that meet all the criteria stated above.

Although the distinction is not totally precise, there are two main approaches, based on:

- Matching the end-user to appropriate assistive technology and/or measuring the outcomes of the use of assistive technology.
- The development of a generic framework for device analysis and synthesis.

The user matching approach has a relationship to the *quality of life* methodologies discussed in Section 1.2. A good example of this approach is the matching person and technology model of Scherer and colleagues (for example, Scherer and Craddock 2002; Scherer and Glueckauf 2005). The model has the following three main components:

- Person, which considers the user's personal characteristics and temperament.
- Milieu, which covers the characteristics of the settings in which assistive technology will be used.
- Technology, which considers the characteristics of the assistive technology, including design factors and funding.

As discussed in more detail in Section 1.2.4 on quality of life assessment, the matching person and technology model is used as part of an assessment process which involves completion of assessment forms. Therefore the end-user is involved in the process of fitting the model. There are a number of other modelling approaches that can be used to match (particular groups of) disabled people to assistive technology, including the following:

- SETT, which is used with disabled students and considers the four areas of student, environment, tasks and tools (Zabala 1998).
- Education Tech Points, which is based around the six 'Tech Points' of referral, evaluation, extended assessment, plan development, implementation and periodic review (Reed and Bowser 1998).

An associated methodology is the Siva Cost Analysis Instrument (SCAI), which is intended to assist the assistive technology professional and the end-user in estimating the costs of a particular assistive technology system and comparing the costs of different options (Andrich 2002).

The USERfit model aims to provide a structured framework for a user-centred approach to assistive technology design (Poulson and Richardson 1998). It consists of three main components: problem definition, functional specification and build and test.

The problem definition component is divided further as follows:

- Context of use, comprising environmental context, which includes a high level summary of the product and its likely users; and product environment, including training, documentation, installation, maintenance and user support for the product.
- Analysis, comprising user analysis, which identifies the stakeholders to be considered in product development and their attributes; activity analysis, covering the range of activities the product will be used for and the associated implications for product design; and product analysis, covering the functional aspects of the product listed as operational features.

Functional specification consists of the following three components:

- Product attribute matrix, which summarises the match between the emerging functional specifications and product attributes.
- Requirements summary, in which the design features identified by the user and activity analysis and their degree of match to user requirements are summarised.
- Design summary in which a more detailed summary of the product functional specifications and their operational details are given.

Build and test consists of usability evaluation, which covers plans for evaluation, objectives, methods and criteria and documentation of the match between the criteria and results.

Human activity assistive technology (HAAT) model

This model due to Cook and Hussey (2002) is an example of the approach based on developing a generic structure which can be used for device analysis, synthesis and development, but not for matching the device to the person. Cook and Hussey start from the definition of an *assistive technology system* as enabling a person to perform an activity in the context of a social environment with the possible aid of some assistive technology. Thus the HAAT model defines an assistive technology system through the following four components:

1. *Context*, which defines the social framework and the physical environment in which the human person and the assistive technology have to operate. The context is divided into the cultural context, the social context, the human setting and the physical context, with further subdivisions suggested for each of these categories.
2. *Human person* represents the human being at the centre of the HAAT model, who is considered to have the attributes of sensory inputs, central processing power and motor inputs. However, this rather mechanistic approach has the disadvantage of reinforcing the tendency of engineers and designers to ignore the wider aspects of human-centred design, such as aesthetics and values, which may be very important in determining whether a device is actually used and whether it actually meets users' needs.

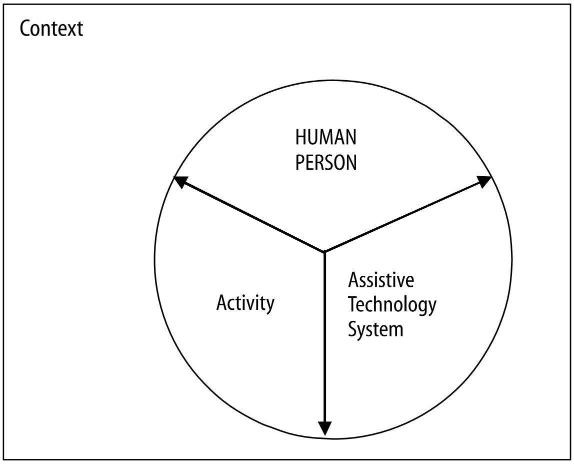


Figure 1.1. The HAAT model of an assistive technology system (after Cook and Hussey 2005)

3. *Activity* defines the procedure, operation or task that the human person wishes to achieve. The activity component is one of the more flexible terms within the model and depends on how the model is being used.
4. *Assistive technology* defines the external enabler used to overcome any contextual barriers or obstacles.

The HAAT model is presented in diagrammatic form in Figure 1.1.

International standards for assistive technology

International standards are relevant to modelling assistive technology. The current international standard which is often found in assistive technology databases and catalogues is ISO 9999: 2002 Technical Aids for Persons with Disabilities, Classification and Terminology. It is product and application orientated and divides assistive technology products into ten classes (with further divisions and subdivisions). The classes include aids for personal mobility, communications, information, signalling and recreation. Whilst the classification is comprehensive, it does not give a basis for understanding the structure of assistive technology systems. It is not a framework that can be used to analyse existing systems or synthesize new systems.

1.3.2 Modelling Human Activities

One aspect of assistive technology modelling that has received a certain, though again limited, amount of attention is that of modelling the human activities for which assistive technology support may be required. There are a number of different approaches depending on the desired end-use of the model and/or the perspective of the modeller. The section will discuss approaches derived from occupational therapy and the EU HEART project.

Table 1.1. Examples of common subtasks

Category	Self-care	Work and school	Play and leisure
Example 1	Communication	Home management	Entertainment
Subtask	Computer access (e-mail)	Computer access (Internet)	Computer access (games)
Example 2	Mobility	Work	Relaxation
Subtask	Moving around a room (home) environment	Moving around a room (office) environment	Moving around a room (bar) environment

The occupational therapy classification uses the Uniform Terminology for Occupational Therapy (Dunn *et al.* 1994, 1995). It is based on three basic performance areas: daily living, work and productive activities, and play and leisure. Whilst these categories are useful for end-user assistive technology system assessment, there is a degree of ambiguity and overlap between them and, for engineering purposes, many of these activities need to be decomposed into a set of more fundamental task components. Table 1.1 shows two examples where there is a common subtask across occupational therapy activities.

The EU HEART project (HEART 1995) defines the four technical areas of communication, mobility, manipulation and orientation. Each area is further divided into activities and then sub-divided into categories of assistive technology devices. The activity categories for the four areas are as follows:

- Communications: interpersonal communications, reading/writing, computer access/user interfacing, telecommunications.
- Mobility: manual mobility, powered mobility, accessibility, private transportation, public transportation, orthoses, prosthetics, seating and positioning.
- Manipulation: environmental control, robotics, orthotics and prosthetics, recreation and sports.
- Orientation: orientation and navigation systems, cognition.

These categories can be further divided into different assistive devices. For instance, interpersonal communication includes hearing aids, optical aids, speech output: recorded and synthetic speech and low-tech devices, such as communication boards. Powered mobility includes powered wheelchairs, powered aids for lifting and transfer and robotic arms for wheelchairs.

There is therefore a need for a new classification scheme of the activity component to avoid ambiguity in task areas and to retain the link with what people actually do in their daily lives. The classification of human activities in the new comprehensive assistive technology model to be outlined in the next section was defined to meet the following specifications:

- To cover all (major) human activity areas whilst minimising overlap between different categories.
- To define activity areas which are sufficiently specific and are neither too wide nor too narrow. This appropriate breadth of categories results in definitions of

activity which are sufficiently precise to be useful, while avoiding an excessively large number of categories.

1.4 The Comprehensive Assistive Technology (CAT) Model

1.4.1 Justification of the Choice of Model

Model development generally involves trade-offs between comprehensiveness and simplicity, since complexity increases with the number of factors and aspects covered. An assistive technology framework which can be used for classification, synthesis and analysis of particular devices, to support the development of new devices and the appropriate matching of technology to the user is likely to be very complicated. This complexity can be managed by the choice of an appropriate modelling framework. The CAT model presented here has been developed out of the HAAT model framework introduced by Cook and Hussey (2002). This approach been chosen for the following reasons:

- It can be represented as a tree-like structure, as illustrated in Figures 1.2–1.6, with a limited number of variables on each branch. This makes the model much easier to understand. It should be noted that the diagrammatic form of the tree structure representation is given solely for the convenience of sighted readers and that the model is not inherently visual. The tree representation gives a way of structuring information to make it easier to understand, use and remember.
- The top level of the model contains the four components which define all assistive technology systems, namely the user, who should be at the centre of assistive technology design; the context(s) in which they will be using the assistive technology; the activities for which they want to use it; and the technology itself.
- The CAT approach gives a generic framework for the categorisation, development, assessment and person matching of assistive technology systems. This framework covers all the main factors, including the social and engineering dimensions, of assistive technology.
- The tree-structure approach is very flexible. As presented the model is comprehensive, without being complicated, due to the tree structure and it can easily be simplified by omitting variables that are not important in the particular context.

The CAT model provides a considerable extension of the approach taken in the HAAT model. The model of the person has been extended to avoid reinforcing the tendency of engineers and designers to ignore the wider aspects of human-centred design, such as aesthetics and values. It is these wider human aspects that often determine whether a device is actually used or rejected and whether it really meets users' needs. Further, a different decomposition of the activity category has been introduced that is based on a set of three fundamental activities and a set of three contextual activities that a person may have to accomplish. The CAT model can be presented in various ways, one of which is in a graphical form; the model is discussed in the following sections.

1.4.2 The Structure of the CAT Model

The comprehensive assistive technology model can be given a tree-like structure that does not use an excessive number of branches at any level. This type of structure has the advantages of being easy to navigate and easy to modify by editing the different steps. It could also be easily translated into an interactive software.

Although the tree structure of the model can most easily be represented diagrammatically, as illustrated in Figures 1.2–1.6, the CAT model is not inherently graphical and visual. It is rather based on decomposed layers of attributes that cover the relevant aspects of a person, their environment and an assistive technology system being used to support their activities. These attributes are ordered to give a structure which makes sense of the different components, and this structure can be represented diagrammatically, as attribute sets or in words, amongst other ways.

The top level of the model has four branches, as illustrated in Figure 1.2 and these branches represent the four components that define all assistive technology systems:

- Context (in which the assistive technology will be used).
- Person (the user, who should be at the centre of assistive technology design).
- Activities (for which the assistive technology will be used).
- Assistive technology.

The model is then developed as an ordered structure by defining further branches at lower levels for each of these top-level categories. The definitions of the branches at the lower levels are motivated by the requirement to use the CAT model in analysis and synthesis (specification development) investigations of assistive technology. The next four sections present the detailed branching structure from the first level branches of Context, Person, Activities, and Assistive Technology.

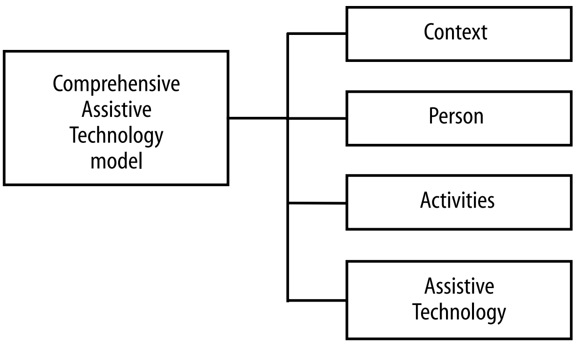


Figure 1.2. Comprehensive assistive technology (CAT) model

1.4.2.1 Context in the CAT Model

This section describes and discusses briefly the second and third levels for the context attribute. The second level describes the main types of context, whereas the third level specifies these contexts in more detail. The full model for the context attribute is given in Figure 1.3. It is important that assistive technology design is based on the user's existing context and does not require them to change it, for instance by requiring a modern infrastructure or the use of English. However, the provision of assistive technology can enable the user to work towards improving their context, either on their own or together with other people.

There are three main types of context at the second level, which can be stated as follows:

- Cultural and social context.
- National context.
- Local settings.

Context attribute – social and cultural component At the third level the cultural and social component is divided as follows (see Figure 1.3):

- Wider social and cultural context.
- User's social and cultural context.

These two contexts may be very similar, particularly in the case of people who form part of the dominant cultural and social context. There are likely to be differences for members of minority groups, including disabled people. Variables of interest in both the user's and the wider social and cultural context include language, other cultural factors, attitudes to disabled people and attitudes to assistive technology. Language and other cultural factors are particularly important since many features of existing assistive technology devices are only provided in English and sometimes a small number of European languages. Both speech output, if any, and documentation and manuals need to be in the local language. In addition, the

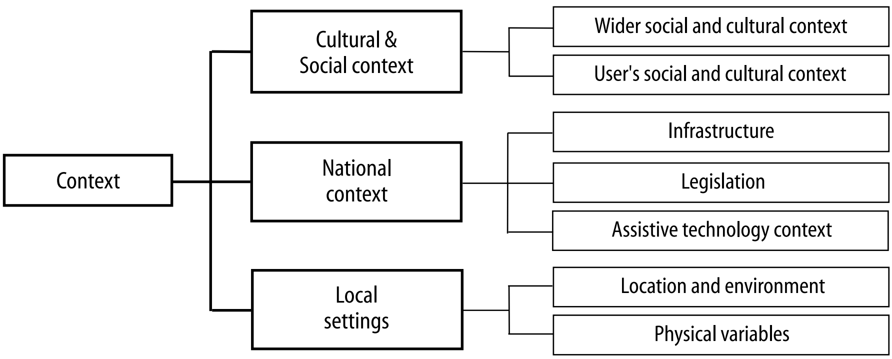


Figure 1.3. Model for context attribute

device needs to be designed and presented in a way that is culturally relevant. This includes the choice of symbols or other labels for controls which are unambiguous in the particular cultural context.

Context attribute – national component The national component is important for distinguishing between the contexts in countries with very different histories, socio-economic situations and states of infrastructure development. Together with the wider social context, this is one of the features of this model that makes it applicable beyond the industrialised countries. The third level of the national component is as follows (see Figure 1.3):

- Infrastructure.
- Legislation.
- Assistive technology context.

The infrastructure context includes factors such as the availability and reliability of an electricity supply and the state of development of information and telecommunications technology, as well as the proportion of the population that have access to them. Other infrastructure factors are the development of a road and rail infrastructure and the extent to which public transport and public and private buildings are accessible, as well as the state of repair of the infrastructure. It would clearly not be very useful to supply an assistive device that requires a mains power supply in a country where the electricity supply is intermittent and unreliable and/or the majority of the population are not connected to the supply.

The legislative context includes the state of legislation about accessibility and social inclusion for disabled people, any relevant building or other regulations and national standards that affect assistive technology. It also includes relevant standards and regulations, as well as any constraints on the use of particular types of assistive technology arising from the legislation, regulations and standards.

The assistive technology context includes the local, regional or national system(s) for distributing, paying and providing support, training and maintenance for assistive technology. This includes the availability of direct payment schemes for disabled people to employ personal assistants or purchase assistive technology of their choice, the extent of availability of sign-language interpreters and guide-communicators for deafblind people and how these services are funded.

Context attribute – local setting component The local settings context describes the various settings in which the user would wish to use assistive devices. This variable has the following two categories at the third level (see Figure 1.3):

- Location and environment.
- Physical variables.

The local and environmental context includes indoors, outdoors or both, by the user on their own or accompanied by other people. Indoor settings could be further classified according to the type of building or its main use, for instance as an apartment in a multistorey block, one storey house, multistorey house or as a home, workplace, educational establishment or hospital. Outdoor settings can be further classified into urban, rural or natural/'wild', *e.g.* mountain or seaside. It should be noted that some devices are intended for use in one particular setting or type of setting, whereas others may be used in several different (types of) settings. For instance, a lift with Braille markings will be used inside in a multistorey building, though the purpose of the building may vary, whereas a wheelchair may be used both inside and outside and in a variety of different types of location and terrain. The local settings context includes any constraints or limiting factors arising from the setting, such as door or room size or the need not to disturb other people.

Physical variables include temperature, noise levels and types, humidity, the level and type of illumination and the types of surface and the state of repair of the local infrastructure. In this case state of repair refers to the specific local environment, whereas in the national context it is more general.

1.4.2.2 Person Attribute in the CAT Model

The person or people who are going to use a particular assistive technology are central to the success of the system. This section discusses the second and third levels of the model for the person attribute. The features of relevance in the person attribute are divided into the following three categories (see Figure 1.4):

- Characteristics.
- Social aspects.
- Attitudes.

Person attribute – characteristics At the third level, the characteristics variables can be divided into the following four categories (see Figure 1.4):

- Personal information.
- Impairments.
- Skills.
- Preferences.

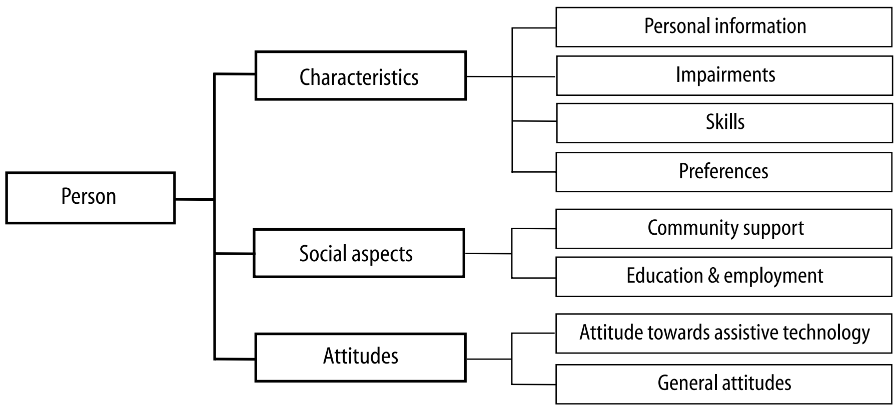


Figure 1.4. Person attribute in CAT model

Personal information includes a range of data, such as age, gender and ethnic origin. Impairments involves the categories of sensory, physical, cognitive, mental health and other impairments. The person's impairments will have an impact both on the activities where they experience barriers and therefore might require assistive technology and the suitable design of this assistive technology so they do not experience problems in using it.

Skills may be innate or the result of education and training. The person will generally have a wide range of very different types of skills, many of which are not particularly relevant to the use of assistive technology. Many disabled people have had bad experiences that have led to a loss of self-confidence and self-esteem. Therefore, it can be of value to look at all the person's skills as part of the process of building self-confidence and self-esteem, which are important for success with assistive technology, before focusing on the more specifically relevant skills. Preferences include preferences for the type of interface, device appearance, receiving basic or detailed information and the way information is presented, for instance as speech, text, pictures or as a combination of several formats.

Person attribute – social aspects At the third level, the social aspects variable can be divided further into the following two variables (see Figure 1.4):

- Community support.
- Education and employment.

Community support may have emotional, practical and/or financial aspects. It involves the availability of support and friendship from the local community and/or family and friends. The availability of support in the form of encouragement can be an important factor in a user having a successful experience with an assistive device, particularly if they experience problems at first. Education and employment include current employment status, employment history, education and training history and qualifications. This variable is likely to be important in a number of ways. Education and employment history can contribute to confidence or lack of it and education and training will have an effect on determining levels of knowledge and skills.

Person attribute – attitudes At the third level the attitudes variable can be categorised as follows (see Figure 1.4):

- Attitudes to assistive technology.
- General attitudes.

User attitudes to assistive technology include user experiences with assistive and other technologies, and how willing users are to try new technology. It also includes whether they prefer to use assistive technology, have a personal assistant, or try to accomplish their goals without either an assistant or the use of assistive technology or to use an assistant for some tasks and assistive technology for others. General attitudes include self-esteem, self-identity, attitudes to disability, (self-)motivation and degree of perseverance.

1.4.2.3 Activities Attribute in the CAT Model

The description of human activities for the purposes of discussing the need for assistive technology to remove accessibility barriers is a complicated issue. A preliminary review was presented in Section 1.3.2. Building on this review, a derivation of the activities component adopted in the CAT model is first presented in this section. This is followed by a full specification of the *activities* component.

1.4.2.3.1 Modelling the Activities Attribute: Specifications

The main aim of the activities component in the CAT model is identifying the accessibility barriers that make it difficult or impossible for disabled individuals and groups to participate fully in all the activities of interest to them. In practice, this requires a classification of human activities, which is equally relevant to disabled and non-disabled people and which covers all the main activities without being so detailed as to be unusable.

A brief overview of existing approaches to modelling human activities was presented in Section 1.3.2. As indicated in this section, the following two main approaches are used to classify human activities in the context of assistive technology:

- Occupational therapy categories based on the following three basic performance areas: daily living, work and productive activities, and play and leisure. This approach is used in the HAAT model.
- Assistive technology products and devices. Although only indirectly a model of human activities rather than assistive technology, this is the basis of the ISO 9999 standard. It is also the approach used in the HEART project after the first level technical area categories.

Both these approaches have their limitations. In particular, both approaches are likely to restrict the range of activities covered to those for which assistive technology categories are available or which are considered appropriate in some sense for disabled people. As a result, some activities, including those where barriers exist for some disabled individuals or groups of disabled people, may be missed. In addition, the occupational therapy category approach is characterised by considerable overlap, as well as some categories that are too fine and detailed and others that are too broad.

Consequently, there is a need for a new systematic classification scheme for the activity component, which includes all the main human activities, without significant overlap and with categories of an appropriate size. Categorising human activities is a complex problem due to the complexity of human behaviour and the wide range of very different types of activities involved. Therefore obtaining a completely decoupled set of categories would require such a fine task decomposition and such a large number of categories or levels of categories as to make the model too large and unwieldy to be of practical use. Therefore, a trade-off will be required between defining the categories sufficiently precisely to avoid overlap and maintaining a reasonable number of categories at each level so that the model remains easy to understand and use.

An associated issue is the type of task decomposition required. Avoiding overlap will give rise to a very large number of simple tasks. However, as well as being unwieldy, this level of task decomposition may be inappropriate for the following two main reasons:

- Many higher-level activities take place within a particular context and have more than one possible decomposition into a sequence of basic tasks. It is generally the ability to achieve the higher-level activity that is important rather than the ability to carry it out in a particular way. Since the contextual information is generally lost when higher level activities are decomposed into basic tasks, this may result in a misunderstanding of what is required and an inability to carry out the higher level activity. For instance, consider the high level activity of environmental control to maintain an appropriate room temperature. This could involve opening a window, adjusting the valve on a radiator, drawing or opening the curtains and/or measuring the temperature of the room. Although measuring the room temperature could be part of this environmental control activity, it is not essential and subjective perceptions of temperature are equally appropriate. In addition, temperature measurement occurs in different contexts, such as environmental control and food preparation. A number of different types of devices can be used to measure temperature, depending on the context. Reading the temperature of the different types of devices may require different tasks or task combinations at a very fine level of decomposition. Consequently, carrying out this fine level of decomposition will lose the connection with the original activity of temperature measurement.
- A very detailed decomposition into basic tasks can lead to overlap between some of the basic tasks resulting from high level activities, which should then be removed. For instance, a number of high level activities, including listening to a radio, choosing the programme on some designs of washing machine and choosing the temperature of an electric oven, involve turning dials. Thus, a model based on decomposition into simple tasks was considered inappropriate.

Taking into account these considerations has led to the following specifications for the model for the classification of human activities in the new comprehensive assistive technology model:

- To cover all (major) human activity areas whilst minimising overlap between categories.
- To have a systematic structure with an underlying logical justification for the categories chosen.
- To define activity areas which are clearly defined and of an appropriate breadth and degree of specificity. The choice of categories with an appropriate breadth results in definitions of activities which are sufficiently precise to be useful, while avoiding an excessively large number of categories.

- To be of real use in identifying the barriers encountered by different groups of disabled people in carrying out activities.
- To be based on a hierarchical structure with two main levels and not more than eight categories on each branch to give a model of appropriate size and complexity to give sufficiently detailed information without becoming excessively large and unwieldy.

1.4.2.3.2 Components in the Activities Attribute

The activities attribute categorises the various activities a person might want to carry out, for some of which they may require support from assistive technology. Following the deliberations of Section 1.4.2.3.1, the logical structure of the categorisation is a division into six activities categories based on the following two main groupings:

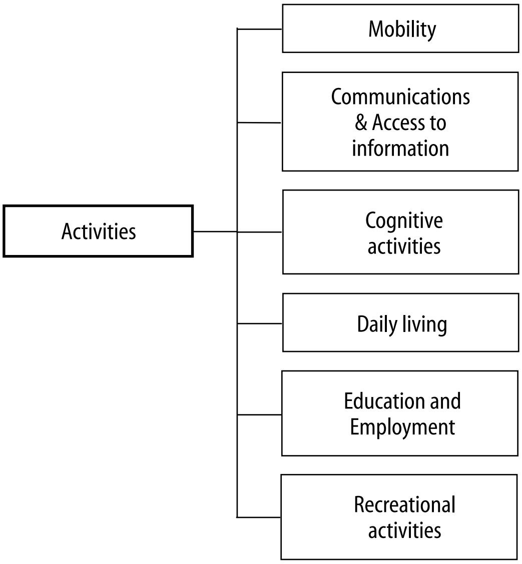
- The major fundamental activity categories of mobility, communications and access to information, and cognitive activities.
- The major contextual activities of daily living, education and employment, and recreational activities.

It should be noted that two of the fundamental activity categories cover most of the activities in the four HEART project technical areas, whereas the contextual activity categories are similar to the occupational therapy performance areas. Thus, the model could be considered to combine the advantages of both of these approaches.

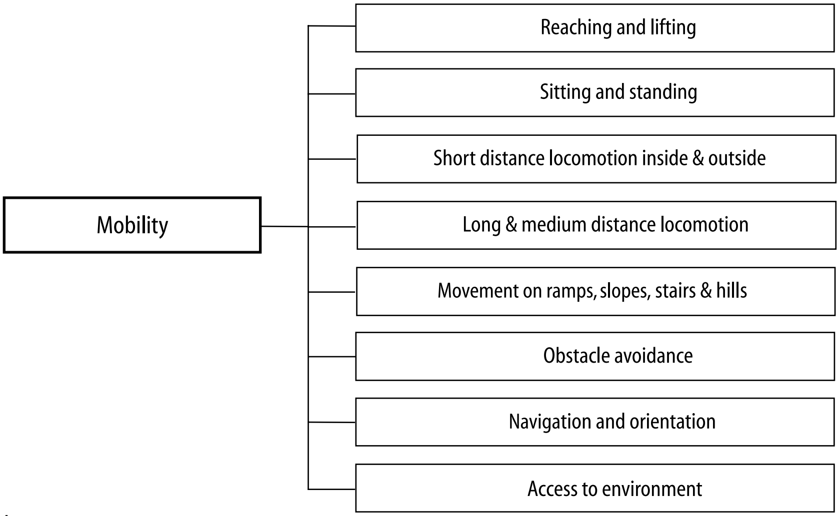
This approach gives the following six main components in the second layer from the activity component of the CAT model (see Figure 1.5):

- Mobility.
- Communication and information.
- Cognitive activities.
- Daily living.
- Education and employment.
- Recreational activities.

It should be noted that each of the contextual activity categories involves activities from all of the fundamental activity categories, namely, some communication and/or use of information, cognitive activities and mobility activities. This is one of the reasons that the activities in these categories are considered 'fundamental'. For instance, computers and the Internet can be used in daily living and the workplace or educational establishment, as well as to obtain information about leisure activities. Many daily living, education and employment and recreational activities require some planning and organising, decision making and analysis of information, as well as sitting and standing and/or lifting and reaching. Therefore, to avoid ambiguity and duplication, cognitive, mobility and communication and information activities are not repeated in the other categories.



a



b

Figure 1.5. a Model for activities attribute. **b** Mobility activities. **c** Communications and accessing information. **d** Cognitive activities. **e** Daily living activities. **f** Education and employment activities. **g** Recreational activities. (*Continued on pp. 30–31*)

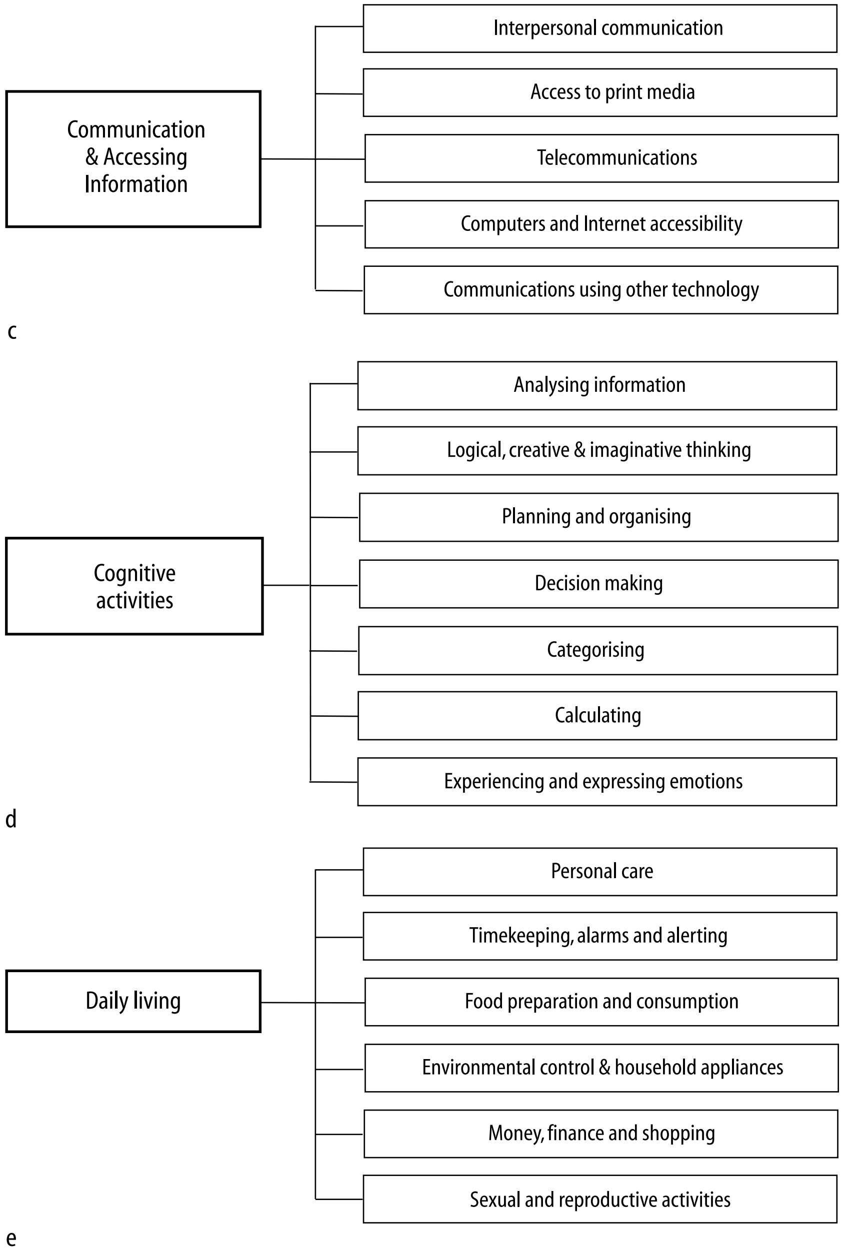
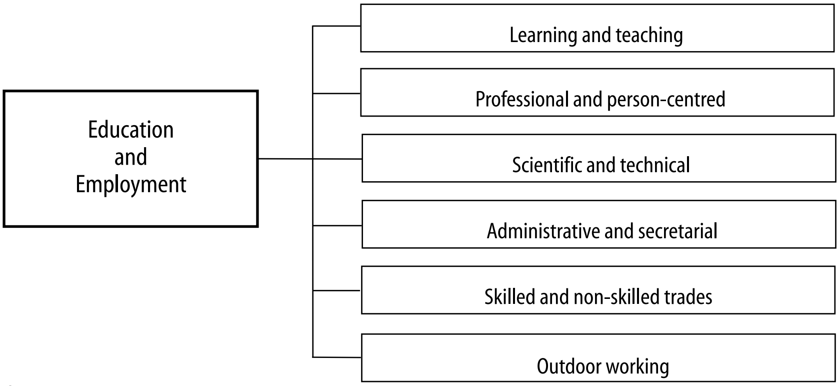
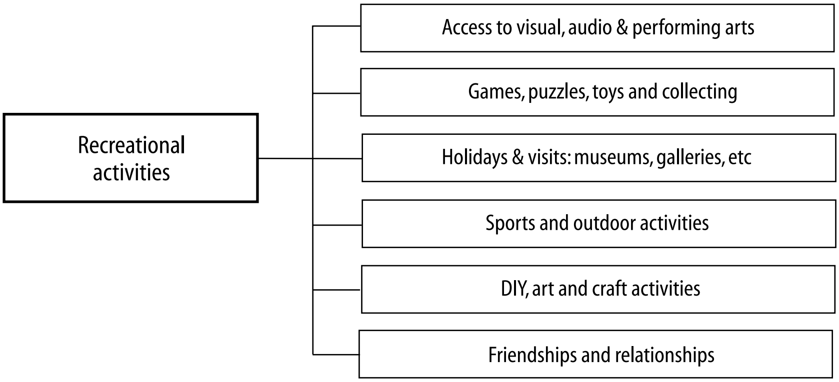


Figure 1.5. (*Continued*)



f



g

Figure 1.5. (*Continued*)

Activities attribute – mobility (Figure 1.5b) Mobility involves all the activities associated with movement and safe travel. This includes avoiding obstacles, as well as the navigation and orientation activities required to reach the right destination by a desired route. It also includes moving objects and sitting and standing. This then gives the following categories at the third level (see Figure 1.5b):

- Reaching and lifting.
- Sitting and standing.
- Short distance locomotion inside and outside.
- Long and medium distance locomotion.
- Movement on ramps, slopes, stairs and hills.
- Obstacle avoidance.
- Navigation and orientation.
- Access to the environment.

It should be noted that long and medium distance locomotion involves travelling either as a passenger or as a driver. In addition, short distance locomotion inside and outside involves issues related to environmental accessibility, *i.e.* the ability to enter, move around and access all the facilities of the location.

Activities attribute – communication and accessing information (Figure 1.5c) Communication and information involves all the activities related to communication, whether carried out directly or *via* technology, as well as those related to accessing information in all its forms. These activities can be categorised further as follows:

- Interpersonal or face-to-face communication.
- Access to print media.
- Telecommunications.
- Computers/Internet.
- Communications and accessing information using other technologies.

Interpersonal or face-to-face communication involves one-to-one, small group and large group (for instance seminars, lectures, meetings and religious services) communication. Access to print media involves both the production of print media and reading print media produced by other people. All the communication categories include artistic and imaginative, as well as factual communications. Communications and accessing information using other technology involves activities using technologies other than print media or information and telecommunications technologies. It includes the use of information kiosks and smart cards.

Activities attribute cognitive activities (Figure 1.5d) Cognitive activities involve all the mental activities or activities related to thought processes. These activities can be categorised further as follows:

- Analysing information.
- Logical, creative and imaginative thinking.
- Planning and organising.
- Decision making.
- Categorising.
- Calculating.
- Experiencing and expressing emotions and feelings.

Logical, creative and imaginative thinking includes dreaming. It should be noted that in general blind and visually impaired people do not require support from assistive technology to carry out cognitive activities.

Activities attribute – daily living (Figure 1.5e) Daily living involves all the different activities used in daily life. Most people will carry out some of the activities in each category on a regular basis and possibly every day or even several times a day. These activities can be further classified as follows:

- Personal care.
- Timekeeping, alarms and alerting.
- Food preparation and consumption.
- Environmental control and household appliances.
- Money, finance and shopping.
- Sexual and reproductive activities.

Personal care activities include personal grooming, washing, toileting and sleeping. They also include personal care activities carried out for a baby or small child. Timekeeping, alarms and alerting includes fire or smoke alarms, alerts to a range of activities, such as the telephone ringing, the door bell (including an indication of who is at the door), a baby crying and a wake-up signal. Food preparation and consumption includes the use of cookers, microwaves and blenders to prepare and cook food. Environmental control and household appliances includes the use of washing machines and vacuum cleaners, as well as turning lights on and off, opening and closing curtains, blinds and doors and control of the heating system. Sexual and reproductive activities include assisted conception, such as donor insemination and in-vitro fertilisation, and the use of contraception. It should be noted that, in general, blind and visually impaired people do not require special assistive technology in this category, other than that required to identify the tablets used in assisted conception and contraception and to access the associated information leaflets.

Activities attribute – education and employment (Figure 1.5f) Education and employment involves the wide range of different activities involved in education, training and employment, both paid and voluntary. Education and employment can further be categorised as follows:

- Learning and teaching.
- Professional and person-centred.
- Scientific and technical.
- Administrative and secretarial.
- Skilled and non-skilled trades.
- Outdoor work.

The scientific and technical activity category includes employment activities in engineering, technology, computing science, programming and information technology, mathematics (other than accountancy and banking) and the natural, physical, social and human sciences. The professional and person-centred activity category covers activities in a wide range of different employment fields, including the health care professions, therapy and counselling, the legal profession, management and personnel, accountancy, banking, religious organisations, translation and interpretation, journalism and creative writing, acting, dancing, music and social and community work.

The administrative and secretarial activity category comprises secretarial and transcription activities, including medical and paralegal transcription, office work, customer services and work in call centres and on information hotlines. Employment activities in the skilled and non-skilled (manual) trades activity category cover the building and decorating trades, shop work, industrial and factory work, janitorial work, the catering and hotel industry, security and monitoring work and the police. Many of the activities in this area involve access to machinery and equipment. Employment in the working outside activity category includes agriculture in all its diversity, gardening, estate management, landscape design, working in zoos and horse riding and racing.

Activities attribute – recreation (Figure 1.5g) Recreational activities can be divided into the following further categories:

- Accessing the visual, audio and performing arts (TV, cinema, theatre, radio, music, dance).
- Games, puzzles, toys and collecting.
- Holidays and visits: museums, galleries, heritage sites.
- Sport and outdoor activities.
- DIY, art and craft activities.
- Friendships and relationships.

Most of the above activities are self-explanatory. The category of craft activities includes drawing and painting, photography, sewing and dressmaking, knitting, crocheting, pottery, metalwork, woodwork and other creative activities or activities that involve making something.

There is a degree of overlap between the three contextual activity categories. For instance, painting and decorating can be either an employment activity (skilled trade) or a recreation activity (DIY and craft activities) and cooking a meal can be either a daily living activity (food preparation) or an employment activity (skilled trade).

1.4.2.4 Assistive Technology Attribute in the CAT Model

Since the community environment is generally not designed for disabled people, they will usually require assistive technology to enable them to carry out many of their desired activities. In some areas, appropriate assistive technology systems are not (yet) available so that some groups of disabled people will have considerable difficulties or even find it impossible to carry out some desired activities. At the second level the assistive technology system attribute can be divided into the following components (see Figure 1.6):

- Activity specification.
- Design issues.
- System technology issues.
- End-user issues.

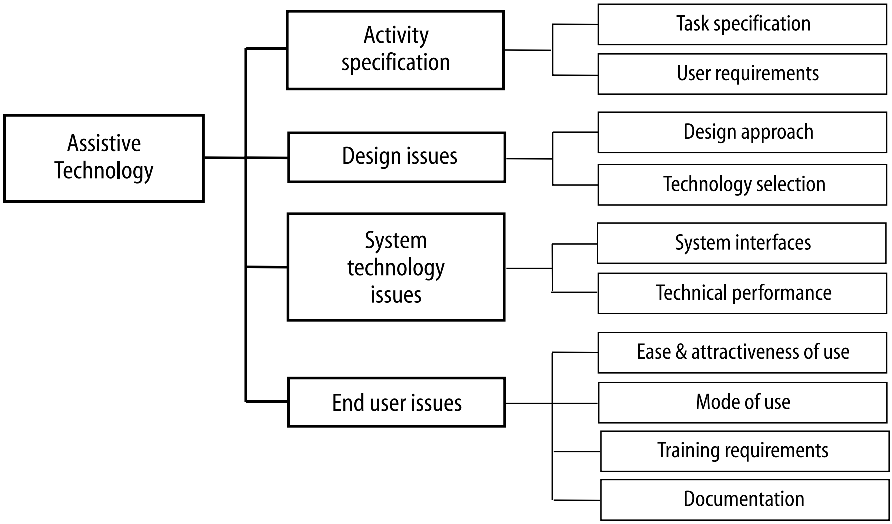


Figure 1.6. Model for assistive technology

System technology issues attribute – activity specification At the third level, the activity specification component can be divided into the following categories (see Figure 1.6):

- Task specifications.
- User requirements.

Task specification involves decomposing a task or activity into a list of the subtasks that the assistive technology has to accomplish. User requirements involve the specification of the physical, sensory and cognitive demands made on the user, such as the ability to lift a certain weight or to follow a sequence of instructions. In general, the fewer and less demanding the user requirements, the wider the group of people who can use a particular technology.

Assistive technology system attribute – design issues At the third level design issues can be categorised as follows (see Figure 1.6):

- Design approach.
- Technology selection.

The design approach is the overall design philosophy or strategy used. This includes whether it is based on design for all or design for specific groups of (disabled) people and/or design for environment or design for maintainability. Design for all aims to design for as wide a population group as possible, independently of factors such as age, gender, ethnic origin, size or disability. Technology selection involves determining the main technologies on which the design could be or is based. For example, in the case of an obstacle detection device the technology choice may lie between infrared and ultrasonic technology and the selection process would consider factors such as performance, the user interface, reliability, technical specifications and cost.

Assistive technology system attribute – system technology issues At the third level, the assistive technology system can be divided into the following further categories (see Figure 1.6):

- System interface.
- Technical performance.

Specification of the system interface includes whether it is single or multimodal and which modalities are available, for instance speech, text, icons and/or Braille. Technical performance includes a range of factors, such as reliability, robustness, safety features and ease of maintenance.

Assistive technology system attribute – end-user issues At the third level, end-user issues can be further categorised as follows (see Figure 1.6):

- Ease and attractiveness of use.
- Mode of use.
- Training requirements.
- Documentation.

Ease and attractiveness of use includes a range of factors, such as whether the assistive technology is designed to be 'user friendly' and intuitive to use, as well as size, weight, portability and appearance. Mode of use includes whether the device is portable or remains in a given location and whether it is stand-alone or forms part of a larger system. Training requirements specify the types of training (different groups of) users are likely to require in order to use the assistive technology successfully. This may involve training before using the technology for the first time and ongoing training. Factors associated with documentation include the formats it is available in, such as on-line and hard copy, how it is structured and whether it is targeted at novice or expert users or there are different sections for both these groups of users.

1.5 Using the Comprehensive Assistive Technology Model

The comprehensive assistive technology model extends the basic structure of the Cook and Hussey HAAT model to encompass a much wider range of attribute components. These are components that the assistive technology engineer may have to consider when constructing or describing a new or an existing assistive technology system. Thus, an important feature of any model is its power in use and in this section some applications of the model are given. First, two of the attribute models are used to give insight into two specific queries:

1. What activity areas should be the focus of assistive technology developments for visually impaired and blind people?

2. Does the CAT model provide further insight into the structure of assistive technology systems?

Following this, a general investigative framework is extracted and used to demonstrate the analysis and synthesis (model specification) capabilities of the CAT model through the presentation of specific case studies.

1.5.1 Using the Activity Attribute of the CAT Model to Determine Gaps in Assistive Technology Provision

One of the aims of the development of the modelling framework is to use it to determine where there are gaps in assistive technology system provision. This may lead to the development of assistive technology systems for areas where there are currently no systems available or lead to developments to extend the capabilities of existing systems to offer more options to end-users.

Many assistive technology systems are developed for disabled people with a particular impairment. In addition, assistive technology aims to overcome the barriers to participation resulting from inaccessible infrastructure and social barriers. Therefore the gaps in assistive technology provision result from the current relationship between assistive technologies, infrastructure, social and other barriers and the disabled person. Consequently, adopting a design for all approach and/or making a particular type of infrastructure accessible would remove the need for assistive technology in that area. For instance, currently many lifts do not have Braille or other tactile markings and audio announcements of the lift operations, such as stopping on level 4. Therefore the provision of Braille or other tactile markings and audio announcements should currently be considered as assistive technology. However, due to the requirements of legislation such as the UK Disability Discrimination Act 1995, it is likely that both Braille and other tactile markings and audio announcements will become standardised in lifts in a number of countries. When this happens, Braille and other tactile markings and audio announcements will no longer need to be considered as assistive technology and this will no longer be an area where blind and visually impaired people require assistive technology to overcome the barriers that would otherwise exist. However, even if design for all becomes standard practice, there are still likely to be areas where assistive technology is required.

As is frequently the case in modelling, the categorisation of activities in Figure 1.5a has required trade-offs between comprehensiveness, that is, a very detailed and all encompassing categorisation and complexity. Within these limits the categorisation has been made as general as possible. In addition, to avoid ambiguity and overlap and to give specificity, in some cases decisions have had to be made as to the appropriate category out of several possibilities for a particular activity. After the initial categorisation, a further type of classification of activities can be carried out:

1. Fully accessible without assistive technology.

2. Require some degree of modification, but not assistive technology, to make them fully accessible.
3. Require assistive technology to make them fully accessible.

Since this book is concerned with assistive technology for blind and visually impaired people, this further classification has been performed for this group of disabled people. It is shown in Table 1.2 with regular text, italics and bold italics for activities considered to be in categories 1 (accessible), 2 (small modifications) and 3 (assistive technology needed) respectively.

1.5.2 Conceptual Structure of Assistive Technology Systems

The CAT model describes assistive technology in terms of the four components of context, person, activity and assistive technology. All these components share interfaces with each other. Thus, focusing on the assistive technology system attribute shows that it has interfaces with the context, person and activity components, as illustrated in Figure 1.1. If the assistive technology system is then given a "processor" core function to co-ordinate the operations between these interfaces and provide the end effector and activity outcomes, a conceptual framework for the assistive technology system emerges. It is possible to build on this revealed structure and use engineering block diagram topology to provide the assistive technology system component with a more formal representation, as is shown in Figure 1.7.

The block diagram model of Figure 1.7 can be applied to the analysis and synthesis of assistive technology.

1.5.3 Investigating Assistive Technology Systems

The second level components of the CAT model attributes of context, person, activity and assistive technology system can be used to develop a checklist of factors to be considered when analysing existing assistive technology systems or developing specifications for new systems. For ease of use these factors are displayed in tabular form in Table 1.3. In some cases it may be necessary to include the third level factors, which can be listed under the associated second level factor.

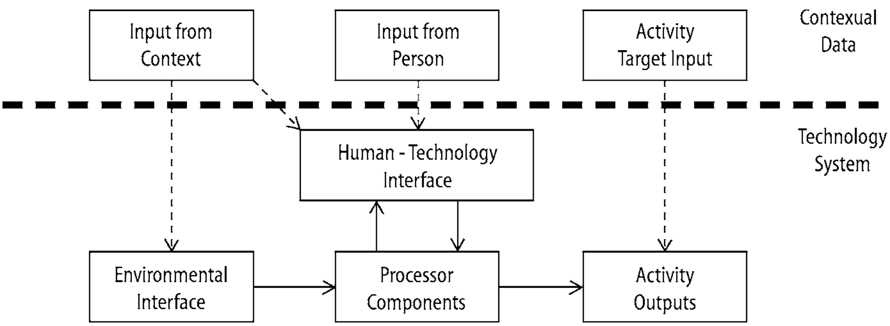


Figure 1.7. Block diagram for assistive technology system

Table 1.2. Issues for accessibility: visually impaired and blind people

	Category of activity	Accessibility status
ACTIVITIES: VISUALLY IMPAIRED AND BLIND PEOPLE	**Mobility**	Reaching and lifting Sitting and standing Short distance locomotion inside and outside Long and medium distance locomotion Movement on ramps, slopes and stairs ***Obstacle Avoidance*** ***Navigation and orientation*** ***Access to environment***
	Communication and access to information	*Interpersonal communications* ***Access to print media*** *Telecommunications* ***Computer and Internet access*** ***Communication using other technology***
	Cognitive activities	Analysing information Logical, creative and imaginative thinking Planning and organising Decision making Categorising Calculating Experiencing and expressing emotions
	Daily living	*Personal care* *Timekeeping, alarms and alerting* *Food preparation and consumption* *Environmental control and using appliances* *Money, finance and shopping* Sexual and reproductive activities
	Education and employment	***Learning and teaching*** ***Professional and person-centred*** ***Scientific and technical*** ***Administrative and secretarial*** ***Skilled and non-skilled*** ***Outdoor working***
	Recreational activities	***Access to visual, audio and performing arts*** *Games, puzzles, toys and collecting* ***Holidays and visits: museums, galleries, heritage sites*** ***Sports and outdoor activities*** ***DIY and craft activities*** *Friendships and relationships*
KEY	Regular text *Italic text* ***Bold italic text***	No accessibility barrier Mild accessibility barrier Severe accessibility barrier

Table 1.3. Assistive technology system – Investigative framework

Attribute	Component	Factors
Context	Cultural and social context	Wider social and cultural issues User's social and cultural context
	National context	Infrastructure Legislation Assistive technology context
	Local settings	Location and environment Physical variables
Person	Social aspects	Community support, Education and employment
	Attitudes	Attitudes to assistive technology General attitudes
	Characteristics	Personal information Impairments Skills Preferences
Activity	Mobility Communication and accessing information Cognitive activities Daily living Education and Employment Recreational activities	This is application specific. The activities of interest in the particular context should be listed.
Assistive Technology	Activity specification	Task specifications User requirements
	Design issues	Design approach Technology selection
	System technology issues	System interface Technical performance
	End-user issues	Ease and attractiveness of use Mode of use Training requirements Documentation

The use of these tools for analysing existing assistive technologies and supporting the development of new assistive technologies are illustrated in the following sections.

1.5.4 Analysis of Assistive Technology Systems

The examples chosen for the analysis investigation are taken from assistive technology systems that have been developed for visually impaired and blind people.

Example 1.1: Ultrasonic cane A visually impaired person is using an ultrasonic cane to avoid obstacles while travelling on foot. Ultrasonic sensors are used to detect obstacles, the signals are processed using signal processing systems and the information is conveyed to the person by vibrating buttons in the handle. An analysis of this assistive technology system using the checklist in Table 1.3 is presented in Table 1.4.

Example 1.2: Accessible lift (elevator) The second example concerns a lift which is used by a wide range of people in a public building, such as a local government office, hospital or further education college. To make the lift accessible to blind people, audio announcements and either Braille or raised (tactile) letters on all control buttons are required. An analysis of a lift which is accessible to blind people is presented in Table 1.5.

Thus, the CAT model for assistive technology can be used to clarify the important features and provide a (technical) definition for existing assistive technology systems. Engineering professionals could use a more detailed version of this decomposition in an assessment of the assistive technology system performance. The analysis of a lift that is accessible to blind people shows a strong dependence on audio-transmitted information; thus this assistive technology solution would not be appropriate for a deafblind person.

1.5.5 Synthesis of Assistive Technology Systems

Another application of the CAT model is in developing specifications for new assistive technology systems.

Example 1.3: Communication for deafblind people This example considers the development of a communication device for a deafblind person who communicates using the British deafblind manual alphabet and reads Braille fluently. A typical scenario might be a brief conversation in the home of a deafblind person, who wants to find out from the postman what has happened to an expected parcel that has not arrived. The application of the checklist in Table 1.3 is given in Table 1.6.

One immediate result of the synthesis exercise is to produce a possible set of communication options that can be investigated further. Figure 1.8 presents some of these options.

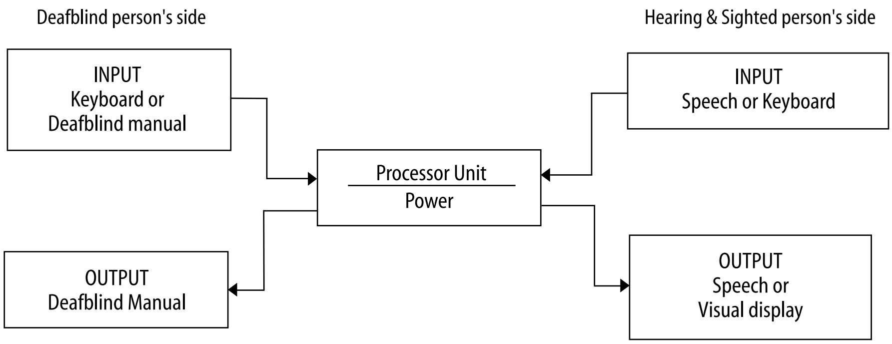


Figure 1.8. Possible communication options in a deafblind personal communicator

Table 1.4. Assistive technology analysis, obstacle avoidance cane using ultrasonic technology

Attribute	Component	Factors
Context	Cultural and social context	Assistive technology is available, but detailed information about available devices is not readily accessible; Travelling alone on public transport and on foot acceptable in the society
	National context	Modern infrastructure. Legislation against disability discrimination
	Local settings	Urban community, noisy indoor and outdoor environment, walking surfaces of varying types and qualities; stationary and moving obstacles
Person	Social aspects	Community support depends on a personal nework of family and friends and some support from organisations; some training for new assistive technology is available
	Attitudes	Willing to try new assistive technology systems and will persevere
	Characteristics	Visually impaired or blind; able to learn and interpret novel sensory information; preference for independent travel
Activity	Mobility	Obstacle location; related to navigation and orientation
Assistive Technology – Ultrasonic Cane	Activity specification	Travelling on familiar and unfamiliar routes in urban environment; avoidance of stationary and moving obstacles in the path of travel
	Design issues	Design for a specific group of disabled people (blind and visually impaired); technology options are laser or ultrasonic sensing technology
	System technology issues	Environmental interface Ultrasonic sensors for obstacle detection: cane body and tips; must be robust to wear and use in different weather conditions
		Processor components Onboard electronics to analyse the incoming ultrasonic signals and drive the tactile (vibrating buttons) user interface; onboard power pack
		Person–technology interface Vibrating buttons in the handle to indicate status of forward path: obstacle free or obstacle present; power level indicators; handle robust to wear, easy to hold and interpret tactile signals; folding cane for storage when not in use
		Activity outputs Information on obstacles in forward path enables safe local travel; fail-safe mode – use as long cane
	End-user issues	Hand-held, lightweight and easily portable, high-tech look; multimodal, audible and tactile alerting to obstacles; Training required for a relatively short period; Documentation to be provided in large print, Braille and on an audio cassette

Table 1.5. Assistive technology analysis, blind-person accessible lift

Attribute	Component	Factors
Context	Cultural and social context	Assistive technology is available, but detailed information about available devices is not readily accessible; Person travelling alone acceptable in the society; Disabled people will accept assistance from strangers
	National context	Modern infrastructure. Legislation for the provision of accessible services for disabled people
	Local settings	Urban community; indoor environments with some noise; office blocks, shopping malls; hotels; people moving in confined space
Person	Social aspects	Workers in the building may be able to provide support; Training in learning Braille and orientation and mobility training available
	Attitudes	Will persevere to achieve objectives
	Characteristics	Blind person with unimpaired hearing or a mild hearing impairment; able to read Braille
Activity	Mobility	Short distance mobility; related to navigation and orientation
Assistive technology – blind-person accessible lift	Activity specification	Finding and travelling in an unfamiliar lift
	Design issues	Design for a specific group of disabled people (blind); Technology options are for tactile and audio technology
	System technology issues	Environmental interface Safety checks on door closing; audio announcements of doors opening and closing; audio announcement of floor number; tactile (raised or Braille) markings on buttons
		Processor components Onboard lift controls accept input from Braille marked buttons; power and speakers for audio system
		Person–technology interface Lift buttons marked in Braille or with raised markings; door status and floor status announced by audio system; audio messages for action in event of lift failure; Braille or raised markings to be robust to wear and tear
		Activity outputs Be able to call the lift; enter the lift; direct the lift to desired floor; leave the lift safely and successfully
	End-user issues	Use of lifts to be part of general mobility training; no documentation requirements for the end-user

Table 1.6. Assistive technology synthesis, communicator for deafblind people

Attribute	Component	Factors
Context	Cultural and social context	Assistive technology is available, but detailed information about available devices is not readily accessible; Person travelling alone acceptable in the society; Disabled people will accept assistance from strangers
	National context	Legislation introduced against disability discrimination
	Local settings	Alone; mainly indoors in home or work environment
Person	Social aspects	Community support depends on a personal network of family and friends and some support from organisations; some training for new assistive technology is available
	Attitudes	Willing to try new assistive technology systems and will persevere; preference for independent living
	Characteristics	Deafblind; able to learn new procedures; uses deafblind manual alphabet and Braille
Activity	Communication	Interpersonal communication
Assistive technology – deafblind person communicator	Activity specification	Face-to-face communication using deafblind manual alphabet with a hearing and sighted person using speech or text
	Design issues	Design for a specific group of disabled people (deafblind people); Technology options need to be determined; only available devices either prototypes or only give one-way communication
	System technology issues	Environmental interface Should be robust, reliable and low maintenance; moderate temperature, dry surroundings, though advantages in water resistance
		Processor components Onboard electronics, signal processing, microprocessor and power pack for portability; two-way communication; receives signal from deafblind manual interface and transforms it to a speech or text signal and outputs the signal on a display or through speaker or headphone for the hearing and sighted person; receives text or speech message signal; transforms it to deafblind manual signal and transmits tactile signal to deafblind person
		Human–technology interface (a) Accepts and generates deafblind manual alphabet signal for deafblind person (b) Accepts and displays text and/or speech for hearing and sighted person Lightweight, portable, robust, intuitive, easy to use
		Activity outputs Interpersonal communication for deafblind person to hearing and sighted person. Simple two-way message transmission
	End-user issues	Hand-held; portable; lightweight; robust; multimodal, audio and tactile interfaces; hearing and sighted person must be able to use the transmission without training; documentation in Braille, deafblind manual alphabet, large print and on audio cassette

This example was intended to be purely illustrative and is not exhaustive of all the potential options. However, a more detailed application of the CAT model could be used to develop detailed design specifications and support the development of new assistive technology systems, in this case a personal communication device for deafblind people.

1.6 Chapter Summary

This chapter considered two main themes as follows.

The social model of disability

There have been some changes in attitudes to and the experiences of disabled people as a result of extensive campaigning by organisations of disabled people, leading to new legislation and some recognition of the need to remove discriminatory barriers and obstacles. In the opening sections of the chapter, the medical and social models for disability were presented and followed by a discussion on quality of life assessment. The role of assessment in prescribing assistive technology solutions was also discussed.

Modelling assistive technology

Social inclusion and the removal of the discriminatory barriers encountered by disabled people require both a design for all approach and the provision of assistive technology systems, where design for all is not feasible. Assistive technology is a complex interdisciplinary area, involving both end-users and a range of different professionals. Therefore the appropriate description and modelling of assistive technology systems is an important issue. The description and modelling of assistive technology systems was the main theme of the chapter and comprised Sections 1.3–1.5. A review of assistive technology systems modelling found only a limited number of approaches and that none of them met the list of specifications drawn up by the authors, giving rise to the need for a new approach. This took as its starting point the human activities assistive technology (HAAT) model. A new CAT model was devised. This new holistic model has a hierarchal tree structure with a number of categories at each level. It can be represented in a number of ways, including as a tree diagram and or a tabular checklist. The tabular checklist model provided a powerful tool for the analysis and synthesis (specification) of assistive technology systems. Examples of analysis and synthesis were presented for three systems of relevance to visually impaired and blind people.

The chapter noted the lack of a suitable categorisation of the *activities* attribute to be used in the CAT model of assistive technology systems. A new general classification of the activities attribute was introduced in the chapter and used to highlight activity areas where visually impaired and blind people require assistive technology systems. This activities categorisation has also been used to structure the book. The first chapters of the book consider the fundamentals of eye physiology, optical parameter measurement and haptics. Subsequent chapters then follow the activity categories of the CAT model, so that the layout of the subjects of the book chapters is as follows.

Fundamentals

Chapter 2 Eye physiology

Chapter 3 Sight measurement

Chapter 4 Haptics and Perception

Mobility

Chapter 5 Overview

Chapter 6 Mobility AT – the Batcane

Chapter 7 Navigation AT – context-aware computing

Chapter 8 Orientation AT – blind person's navigator

Chapter 9 Electronic Travel Aids – an assessment

Chapter 10 Accessible Environments

Chapter 11 Accessible bus system – a Bluetooth case study

Communication and access to information

Chapter 12 Information Technology – a survey

Chapter 13 Magnifiers, screen readers, self-voicing applications

Chapter 14 Speech, text, Braille conversion technology

Chapter 15 Accessing books and documents

Chapter 16 Music for the print impaired

Daily living

Chapter 17 Assistive technology for daily living

Education and employment and recreational activities

Chapter 18 Assistive technology for education, employment and recreation

Questions

Q.1 Compare and contrast the medical and social models of disability.

Q.2 Describe three different approaches to quality of life indices and discuss their relevance (if any) to the assessment of the benefits of assistive technology systems.

Q.3 Select an example of an assistive technology system and identify the key components of the system. Explain how the design is tailored to a specific end-user group.

Q.4 Briefly describe the comprehensive assistive technology model. List and explain briefly the different ways it can be used to investigate assistive technology systems.

Q.5 Use the assistive technology block diagram in Figure 1.7 to describe two different assistive technology systems.

Q.6 Describe the "Design for All" philosophy. What are its advantages and disadvantages?

Projects

P.1 The political organisation of disabled people and the development of a Movement of Disabled People are relatively recent. Use the Internet and other sources of information to list the current organisations that could be considered part of this movement. Summarise the aims of these organisations. Make an assessment of the potential impact of their activities on the following:

1. Political institutions
2. The built environment
3. The development of assistive technology systems
4. The delivery systems for assistive technology

P.2 Use the CAT model and the assistive technology block diagram to analyse the following as assistive technology systems:

1. A guide dog
2. A long cane
3. A laser cane
4. A liquid level sensor
5. Braille playing cards
6. A light probe

P.3 Consider the problem of designing an object detection device for blind and deafblind people. Use the CAT model to obtain the basic description and a list of specifications for such a device.

P.4 Download the sample MPT forms from the Internet. Role play filling in the forms with a colleague. You can either define an appropriate scenario or use information based on someone you know. Analyse the procedure and the content of the forms. Comment on the advantages and disadvantages of the process and the resulting information. Do you think the process is useful in choosing a particular assistive technology? Do you think the procedure empowers or disempowers the disabled person? Do you have any other comments?

P.5 (a) Use the activities diagram to highlight the activity areas where significant barriers are likely to be encountered by:

1. A profoundly deaf person
2. A severely visually impaired person who has tunnel vision and can read

(b) Use the CAT model components to compare and contrast what the model reveals about appropriate provision of an obstacle avoidance technology for a blind and mobile person in a country in Western Europe and a country in Sub-Saharan Africa.

(c) Use the CAT model:

1. To specify an accessible lift technology which can be used by a blind wheelchair user.
2. To investigate the provision of accessible computers for a sighted profoundly deaf person whose first language is a sign language. The accessible computer is to be sited and used in a public library facility which provides access to the Internet and general computing facilities, including word processing and spreadsheets.

References

Andrich, R., 2002, The SCAI instrument: measuring costs of individual assistive technology programmes, Technology and Disability, Vol. 14, pp. 95–99

Barnes, C., 1994, Disabled People in Britain and Discrimination: a case for anti-discrimination legislation, Hurst & Co., London

Beckie, T.M. and Hayduk, L.A., 1997, Measuring quality of life, Social Indicators Research, Vol. 42, pp. 21–39

Bowling, A. 1995, Measuring Disease: A Review of disease Specific Quality of Life Measurement Scales, Open University Press, Buckingham, UK, pp. 1–16

Bowling, A. and Windsor, J., 1999, Towards the good life: a population survey of dimensions of quality of life, J. Happiness Studies, Vol. 2, pp. 55–81

Brisenden, S., 1986, Independent living and the medical model of disability, Disability, Handicap and Society, Vol. 1, No. 2, pp. 173–178

Browne, J.P., O'Boyle, C.A., McGee, H.M., Joyce, C.R., McDonald, N.J., O'Malley, K., and Hiltbrunner, B., 1994, Individual quality of life in the healthy elderly, Qual. Life Res., Vol. 3, pp. 235–44

Calman, K.C., 1984, Quality of life in cancer patients - an hypothesis, J. Med. Ethics, Vol. 10, No. 3, pp. 124–127

Cook, A.M. and Hussey, S.M., 2002, Assistive Technology: Principles and Practice, 2nd Edition, Mosby Inc., St. Louis, USA

Damodaran, L., 1996, User involvement in the systems design process- a practical guide for users, Behaviour and Information Technology, Vol. 15, No. 6, pp. 363–377

Day, H. and Jutai, J., 1996, Measuring the psychosocial impact of assistive devices; the PIADS. Canadian J. Rehabilitation, Vol. 9, No. 2, pp. 159–168

Day, H., Jutai, J. and Campbell, K.A., 2002, Development of a scale to measure the psychosocial impact of assistive devices: lessons learned and the road ahead, Disability and Rehabilitation, Vol. 24, No. 1/2/3, pp. 31–37

Demers, L., Wessels, R.D., Weiss-Lambrou, R., Ska, B. and de Whitte, L.P., 1999a, An international content validation of the Quebec user evaluation of satisfaction with assistive technology (QUEST), Occupational Therapy International, Vol. 6, No. 3, pp. 159–175

Demers, L., Ska, B, Giroux, F and Weiss-Lambrou, R., 1999b, Stability and reproducibility of the Quebec user evaluation of satisfaction with assistive technology (QUEST), J. Rehabilitation Outcomes Measurement, Vol. 3, No. 4, pp. 42–52

Demers, L., Weiss-Lambrou, R. and Ska, B., 2000, Item analysis of the Quebec user evaluation of satisfaction with assistive technology (QUEST), Assistive Technology, Vol. 12, pp. 96–105

Demers, L., Weiss-Lambrou, R and Ska, B., 2002a, The Quebec user evaluation of satisfaction with assistive technology (QUEST 2.0): an overview and recent progress, Technology and Disability, Vol. 14, pp. 101–105

Demers, L., Monette, M., Lapierre, Y., Arnold, D.L. and Wolfson, C., 2002b, Reliability, validity and applicability of the Quebec user evaluation of satisfaction with assistive technology (QUEST 2.0) for adults with multiple sclerosis, Disability and Rehabilitation, Vol. 24, No. 1–3, pp. 21–30

DeRuyter, F., 1997, The importance of outcome measures for assistive technology service delivery systems, Technology and Disability, Vol. 6, pp. 89–104.

Diener, E., 1984, Subjective wellbeing, Psychological Bulletin, Vol. 95, pp. 542–575

Doward, L.C. and McKenna, S.P., 2004, Defining patient-reported outcomes. Value in Health, Vol. 7, Supplement 1, pp. S4–S8

Dunn W., Foto M., Hinojosa J., Schell B., Thomson L.K. and Hertfelder S., 1994, Uniform terminology for occupational-therapy - 3rd Edn. Am. J. of Occup. Therapy, Vol. 48, No. 11, pp. 1047–1054

Dunn W., Foto, M., Hinojosa J., Schell B., Thomson L.K. and Hertfelder, S., 1995, Uniform terminology for occupational therapy - 3rd Edn., Am. J. Occup. Therapy, Vol. 49, No. 4, p. 371

Eid, M. and Diener, E., 2004, Global judgments of subjective wellbeing: situational variability and long-term stability, Social Indicators Research, Vol. 65, pp. 245–277

Fouygerollas, P., Noreau, L., Bergeron, H., Cloutier, R., Dion, S.A. and St-Michel, G., 1998, Social consequences of long-term impairments and disabilities: conceptual approach and assessment of handicap. Int. J. Rehabil. Res. Vol. 21, No. 2, pp. 127–142

Fuhrer, M.J., 1996, The subjective wellbeing of people with spinal cord injury: relationships to impairment, disability and handicap, Topics in Spinal Cord Injury Rehabilitation, Vol. 75, pp. 633–638

Fuhrer, M.J., 2000, Subjectifying quality of life as a medical rehabilitation outcome, Disability and Rehabilitation, Vol. 22, No. 11, pp. 481–489

Fuhrer, M.J., Rintala, D.H., Hart, K.A., Clearman, R. and Young, M.E., 1992, Relationship of life satisfaction to impairment, disability and handicap among persons with spinal cord injury living in the community, Archives of Physical Medicine and Rehabilitation, Vol. 73, pp. 552–557

Fuhrer, M.J., Jutai, J.W., Scherer, M.J. and DeRuyter, F., 2003, A framework for the conceptual modelling of assistive technology device outcomes, Disability and Rehabilitation, Vol. 25, No. 22, pp. 1243–1251

Gelderblom, G.J. and de Witte, L.P., 2002, The assessment of assistive technology, outcomes, effects and costs, Technology and Disability, Vol. 14, pp. 91–94

Gill, T.M. and Feinstein, A.R., 1997, A critical appraisal of the problem of quality of life in medicine, J. American Medical Association, Vol. 278, pp. 619–626

HEART, 1995, Line C Rehabilitation technology service delivery systems in Europe, Brussels, European Commission

Hersh, M.A., 2006, Mathematical Modelling for Sustainable Development, Springer Verlag, Berlin, Germany

Hickey, A.M., Bury, G., O'Boyle, C.A., Bradley, F., O'Kelly, F.D., and Shannon, W., 1996, A new short form individual quality of life measure (SEIQoL-DW): application in a cohort of individuals with HIV/AIDS, BMJ, Vol. 313, pp. 29–33

Hughes, C., Heang, B., Kim, J. Eisenman, L.T. and Killian, D.J., 1995, Quality of life in applied research: a review and analysis of empirical measures, Am. J. on Mental Retardation. Vol. 99, No. 6, pp. 623–641

Jutai, J. and Day, H., 2002, Psychosocial impact of assistive devices scale (PIADS), Technology and Disability, Vol. 14, pp. 107–111

Kim, J.O. and Curry, J., 1977, The treatment of missing data in multivariable analysis, Social. Meth. Res., Vol. 6, pp. 215–240.

Maslow, A., 1968, Toward a Psychology of Being, 2nd Edition, Van Nostrand, New York

McDowell, I., and Newell, C., 1996, Measuring Health: A guide to rating scales and questionnaires, 2nd Ed., Oxford University Press, New York

Mountain, L.A., Campbell, S.E., Seymour, D.G., Primrose, W.R. and Whyte, M.I., 2004, Assessment of individual quality of life using the SEIQoL-DW in older Patients, Q. J. Med. Vol. 97, pp. 519–524

Noreau, L., Fougeyrollas, P. and Vincent, C., 2002, The LIFE-H: assessment of the quality of social participation, Technology and Disability, Vol. 14, pp. 113–118

O'Boyle, C.A., McGee, H., Hickey, A., O'Malley, K., Joyce, C.R.B., 1992, Individual quality of life in patients undergoing hip replacement Lancet, vol. 339, pp. 1088–91.

O'Boyle C.A., McGee, H., Hickey, A., Joyce, C.R.B., Brown, J., O'Malley, K., 1993, The schedule for the evaluation of individual quality of life, User manual, Department of Psychology, Royal College of Surgeons in Ireland, Dublin

Phillips, B., and Zhao, H., 1993, Predictors of assistive technology abandonment, Assistive Technology, Vol. 5, No. 1, pp. 36–45

Poulson, D., and Richardson, S, 1998, USERfit – a framework for user-centred design in assistive technology, Technology and Disability, Vol. 9, pp. 163–171

Reed, P., and Bowser, G., 1998, Education tech points: a framework for assistive technology planning and systems change in schools, CSUN Conference on Technology and Persons with Disabilities, Los Angeles, CA.

Rowley, J., 1998, Towards a methodology for the design of multimedia public access interfaces, J. of Information Science, Vol. 24, No. 3, pp. 155–166

Sawicki, D.S., 2002, Improving community indicator systems: injecting more social science into the folk movement, Planning Theory and Practice, Vol. 3, No. 1, pp. 13–32

Schalock, R.L., 1996, Reconsidering the conceptualization and measurement of quality of life, In R. L. Schalock, and G. N. Siperstein (Eds.), Quality Of Life Volume I: Conceptualization and Measurement pp. 123–139, American Association on Mental Retardation, Washington, DC

Scherer, M.J., 2000. Living in the State of Stuck: How Technology Impacts the Lives of People with Disabilities, 3rd. Edition, Broodline Books, Cambridge, MA, USA

Scherer, M.J., and Craddock, G., 2002, Matching Person and Technology (MPT) assessment process, Technology and Disability, Vol. 14, No. 3, pp. 125–131

Scherer, M.J., and Cushman, L., 2000, Predicting satisfaction with assistive technology for a sample of adults with new spinal cord injuries, Psychological Reports, Vol. 87, pp. 981–987

Scherer, M.J., and Glueckauf, R., 2005, Assessing the benefits of assistive technologies for activities and participation, Rehabilitation Psychology, Vol. 50, No. 2, 132–141

Simon, S.E., and Patrick, A., 1997, Understanding and assessing consumer satisfaction in rehabilitation, J. Rehabilitation Outcomes Measurement, Vol. 1, pp. 1–14

Slevin, M.L., Plant, H., Lynch, D., Drinkwater, J., and Gregory, W.M., 1988, Who should measure quality of life, the doctor or the patient?, British J. of Cancer, Vol. 57, pp. 109–112

Smith, R.O., 1996, Measuring the outcomes of assistive technology: challenge and innovation, Assistive Technology, Vol. 8, pp. 71–81

Smith, R.O., 2002, OTFACT: Multilevel performance-oriented software with an assistive technology outcomes assessment protocol, Technology and Disability, Vol. 14, pp. 133–139

Spangers, M.A.G., and Aaronson, N.K., 1992, The role of health care providers and significant others in evaluating the quality of life of patients which chronic disease: a review. J. Clin. Epidemiol., Vol. 45, pp. 745–760

Swain, J., French, S., and Cameron, C., 2003, Controversial Issues in a Disabling Society, Open University Press, Buckingham, UK

UPIAS, 1976, Fundamental Principles of Disability, London: UPIAS (Union of the Physically Impaired Against Segregation)

Wessels, R., Persson, J., Lorentsen, O., Andrich, R., Farrario, M., Oortwijn, W., VanBeekum, T., Brodin, H., and de Witte, L., 2002, IPPA; Individually prioritised problem assessment, Technology and Disability, Vol. 14, pp. 141–145

WHO, 1980, International Classification of Impairments, Disabilities and Handicaps, World Health Organisation, Geneva, Switzerland

WHO, 2001, International Classification of Functioning, Disability and Health, World Health Organisation, Geneva, Switzerland

WWW1, 2006, http://www.qolid.org/

Zabala, J., 1998, Ready, SETT, Go! Online Workshop,
http://www2.edc.org/NCIP/Workshops/sett/index.html

2 Perception, the Eye and Assistive Technology Issues

Learning Objectives

The sense of sight results from the detection and processing of visual information in the eye, and the visual areas of the brain to which it is transmitted in the form of electrical signals by the optic nerve. One of the distinctive aspects of the sense of sight is the ability to use it to obtain an overview of a situation or scene. This involves the processing of very complex information and the formation of visual maps in the brain. Other features of the sense of sight include the ability to understand and interpret both moving and still images and to perceive both two and three-dimensional objects. Unfortunately the mechanisms of visual processing in the brain and the detailed functions of its different visual areas are still not fully understood.

The chapter comprises five sections, followed by a chapter summary and projects. The opening section, Section 2.1, introduces an overview of perception and seeks to place the sense of sight in this context. A description of the physiology of the visual system then follows in the next section. After the presentation of the main structures of the eye, the focus turns to the retina and the subsequent processing by the nervous system and the brain. Section 2.4, entitled Vision in Action, discusses how the eye accomplishes operations such as image formation, accommodation to focus on objects at different distances and colour vision. Section 2.5 discusses visual impairment and assistive technology design, including the demographics of visual impairment, the main different types of visual impairments and the use of corrective lenses.

The main aim of the chapter is to give an understanding of the main components and functions of the visual system, as well as some of the different types of visual impairments. Specific learning objectives include the following:

- Understanding the context of vision in human perception.
- Knowledge of the main structures of the eye and their functions.
- A basic overview of the visual processing system in the brain.
- Understanding of the activities carried out by the visual system, including spatial form processing, binocular vision, motion detection and colour vision.

- Knowledge of some of the different types of visual impairments and the use of corrective lenses for the common impairments in focussing, such as myopia, hyperopia and astigmatism.

2.1 Perception

2.1.1 Introduction

Perception is the process of obtaining, selecting, interpreting and organising sensory information. In addition to its other functions, this sensory information can be used to support motor activity. Perception can be divided into the following two categories:

- Exteroception or the perception of information from external stimuli.
- Interoception or the perception of information from stimuli within the body. This includes proprioception or awareness of changes in body position and movement of the body and the vestibular system that detects movement and changes in the position of the head.

It is mainly the exteroceptive and proprioceptive systems that are relevant to assistive technology, although there may be a role for the vestibular system. Thus, interoceptive systems are not discussed further here.

The extroreceptive systems can be categorised as:

- Distant senses: vision, hearing and olfaction (smell).
- Contact (or near) senses: tactile system and gustation (taste).

The distinction between distant and contact senses is very important in the context of both sensory processing and assistive technology. It affects the search strategies that can be used to obtain environmental information and the type and extent of information that can be obtained. In addition the distant senses are, in principle, able to obtain an overview of a scene, whereas the contact senses can only obtain information from the part of the scene with which contact can be made. Gustation is probably inappropriate for use in assistive devices, due to the social unacceptability of investigating objects by sticking out the tongue and the possibility of damage to it or unfortunate reactions if objects are brought into the mouth. The strength of olfaction and the ability to use it to identify objects are very variable between different people and olfaction is currently not used in any assistive devices. This leaves hearing and touch as the main senses that blind people can use to obtain external sensory information to complement their sense of sight.

Both hearing and sight are able to process complex information obtained from a wide spatial area and process several items of information simultaneously. However, there are also significant differences. In particular, auditory information is generally sequential, whereas visual information can be either sequential or simultaneous. The ability of the visual system to process simultaneously a large quantity of spatially organised information makes it particularly useful for consulting lists and obtaining an overview of a scene. Therefore, the cognitive load on memory is less when using the visual than the auditory sense.

Perception of the world is inherently multisensory with integration of information from the different senses (Morein-Zamir *et al.* 2003). The different ways in which the different senses function enable them to obtain different types of information, thereby complementing each other. However, there is also some redundancy in sensory information, which can be helpful to people with sensory impairments.

2.1.2 Common Laws and Properties of the Different Senses

Despite the great differences between the senses, there are also common factors. In particular, the smallest change in stimulation (ΔI) which is perceptible is proportional to the magnitude of the original stimulus (I_o), with a constant of proportionality $k < 1$, that is:

$$\Delta I = kI_o$$

This expression is known as Weber's law. It holds for the type of variables known as prothetic continua, where a larger numerical value indicates an increase in quantity, such as pressure or amplitude, rather than quality, such as the frequency of a tone. There is some disagreement as to whether the perceived impact of a signal increases with the logarithm or power, with the logarithmic and power relationships referred to as Fechner's Law and Stevens' Power Law respectively:

$$\text{Fechner's Law: } S = \ log I$$
$$\text{Stevens' Power Law: } S = k_1 I^n$$

where S and I are the perceived impact and the signal respectively and k_1 is a constant of proportionality.

In either case, the sensation increases less rapidly than the stimulus, so that relatively large changes are required at high stimulus intensity before a change is noticeable.

The sensation produced by a stimulus of constant magnitude has two stages, which unfortunately do not correspond exactly to the transient and steady states. In the first stage, the response increases over a time period of anything from a fraction of a second to several seconds, whereas in the second stage it decreases to obtain its final asymptotic value, which may be zero. However, the sensation may be sustained at its highest value for a period before it starts to decrease. Most sensory stimuli exhibit temporal summation or integration, with the magnitude of the sensation increasing with stimulus duration for short time periods.

Time dependent changes in the sensitivity of a sensory system (over slightly longer time periods) are called sensory adaptation. This includes light and dark adaptation, which will be discussed in Section 2.4.3. Both spatial summation and spatial inhibition occur. In spatial summation, the sensation magnitude depends on the area of the sensory surface being stimulated, whether this is small as in the case of the retina or the basilar membrane of the ear or large, as in the case of the skin. However, the details of the summation process, which is not always linear, vary for the different senses and with the stimulus intensity. Brightness contrast is particularly important in vision and remains reasonably constant despite changes in illumination. However, the apparent brightness of a weak visual stimulus is reduced significantly by a surrounding stronger stimulus or inhibitory field, for example, a very bright light reduces the apparent brightness of nearby weaker lights. Auditory masking is the suppression of sounds by louder sounds and loudness under masking is analogous to brightness under contrast. However, despite the similarities between auditory masking and brightness contrast, there are also differences. In particular, there is no real equivalent to total darkness, and loudnesses add, so that a tone plus noise produces a louder sound, whereas spatially separated bright signals generally do not combine with each other to produce a brighter signal. Tactile vibration seems to follow similar masking rules to those of sounds.

2.1.3 Multisensory Perception

There is some evidence that the different senses influence each other. However, additional research will be required to determine whether multisensory integration could be used as the basis of new assistive technology devices. This includes the following possibilities:

- The use of an audio or a tactile signal to strengthen a weak visual signal.
- The use of audio or tactile signals to give rise to visual images.

Some mobility devices or prototypes convey distance and directional information about obstacles in terms of the amplitude and pitch of an audio signal. When this information is perceived purely as auditory information, it may not be particularly easy or natural to interpret. However, if it can give rise to some type of visual perception, this may overcome this problem and make the information easier and more natural to interpret.

Multisensory effects, which will be discussed in this section, include the McGurk effect, the ventriloquism effect, synthaesesia and cross-modal matching. The existence of these multisensory effects indicates that sensory information from different sensory modalities is integrated, rather than separate visual, auditory and tactile worlds being perceived. Both the ventriloquism and McGurk effects are examples of perceptions from one sense influencing judgements about perceptions from another sense, whereas synthesia is the experience of sensory perception from more than one sense when sensory information is supplied to only one sense.

In the McGurk effect a visual stimulus produces an auditory illusion by influencing what is heard. This coordination is used in lip reading. In the ventriloquism effect, the perceived location of a sound is shifted towards a visual stimulus at a different position when the stimuli occur within about 300 ms of each other. This is used to give the impression that a puppet is talking, as auditory perceptions are matched to the visual perception of the movement of the puppet's lips.

Synthaesthia (from the Greek *syn* together and *aesthesis* perception) is the experience of sensory perception from one sense due to stimulation of another sensory modality. Synthaesthia can be described in terms of two interrelated components and involves a triggering inducer, such as a sound, producing a synaesthetic concurrent, such as a colour. It generally involves two rather than multiple senses and can sometimes involve only one sensory modality, as in the case of visual letters giving rise to perceptions of visual colours. It is generally mono- rather than bidirectional. For instance sight is experienced as touch, without touch also triggering visual experiences or *vice versa* (Grossenbacher and Lovelace 2001). Unidirectional synthaesesia is generally pleasurable, whereas bidirectional synaesthesia can cause dizziness and stress. Statistics vary from 1 in 25,000 to 1 in 2000 people having synthaesthia, with it being 3 to 8 times more common in women. Synaethesia is a passive involuntary experience. The parallel sensations often lead to very good memory, particularly of the secondary perception and often of the details of conversations, instructions and the spatial locations of objects.

Cross-modal matching involves the use of information from one sensory modality to make judgements about an equivalent stimulus from another modality. For instance, haptic cues can be used to judge which of a number of objects is most similar to an object that has been seen. Most non-disabled adults and children are able to do this without difficulty. Cross-modal matching has also been carried out to match the intensities of different types of signals, for instance the force on a handgrip to the loudness of a 1000-Hz tone or the loudness of a tone to the brightness of a light. Subjects have been found to perform this type of matching without difficulty. A number of different explanations have been suggested for cross-modal matching, including the conversion of all information to one sensory modality and amodal coding of information.

In multisensory perception, one sense generally dominates or captures the other(s) (Soto-Faraco *et al.* 2004a,b). According to the modality-appropriateness account of sensory dominance, perception will be dominated by the sense that provides the most accurate or appropriate information about the particular stimulus (Welch and Warren 1986). For sighted people this is most frequently vision, unless there are very significant differences in the intensities of the different stimuli. This raises the question of whether hearing or touch dominates for blind people. It is accepted that visual cortex activity generally accompanies tactile perception (Zhang *et al.* 2004). However, this could mean that both visual and tactile processing involve multisensory representations or that the visual cortex is not purely visual. An example of the effect of visual stimuli on the other senses is the special effects in certain films, which induce vestibular and visceral sensations by mimicking the visual cues in, say, a bush plane flying over a mountain range or a roller-coaster ride. The visual perceptions of pilots and astronauts are affected by gravity, particularly during take-offs, landings and space flights. For instance, in a zero gravity environment objects appear to be lower than they are, as the eyes shift upwards.

The integration of auditory and visual stimuli occurs for both meaningful speech and meaningless combinations such as tone bursts and flashing light spots. Reactions are faster and signal localisation is improved for auditory and visual signals that are coincident in time and space compared to separated or unimodal signals. In some cases, the thresholds for sound stimuli are reduced by simultaneous visual stimuli and simultaneous auditory stimuli increase the perceived intensity of visual stimuli. Brief sound bursts and light flashes presented soon after each other are perceived to be closer together in time than they are in actual fact (Lewald and Guski 2003).

Reactions to auditory stimuli are generally about 20–60 ms faster than those to visual stimuli. This is due to the faster transduction in the inner ear compared to that of the retina. However, auditory latencies of the multisensory neurons in the superior temporal sulcus are 95 ms shorter than visual latencies. Consequently, presenting auditory stimuli 50–100 ms after visual stimuli gives the optimum circumstances for them to be perceived as simultaneous (Lewald and Guski 2003). On the other hand, the speed of sound is only a small fraction of that of light (330 rather than 299,792,458 m/s), so that acoustic stimuli will lag visual ones when they are at a distance from the observer. This is why lightning is seen before the associated thunder is heard. The differences in processing time and travel speed cancel each other out for visual and auditory stimuli at a distance of about 10 m, which is called the horizon of simultaneity. However, audiovisual events at other distances are perceived as synchronous and a lag of more than 250 ms is required before the asynchronicity between auditory and visual speech cues is perceived. This perceived synchronicity is important for lip-reading. It has been suggested that insensitivity to small differences in the arrival times of stimuli in different sensory modalities is due to the window for multisensory temporal integration increasing with the distance from the observer up to a distance of about 10 m. This is equivalent to a temporal ventriloquism effect and the stimuli are perceived as simultaneous for distances up to 10 m, which is the horizon of simultaneity (Spence and Squire 2003).

The combination of stimuli generally gives an increased response over a single sensory stimulus in all multisensory categories. Spatially coincident multisensory stimuli generally produce response enhancement, whereas spatially separated stimuli may either reduce the response or not affect it. There is also a long temporal window during which stimuli can affect each other. This temporal windows takes account of the fact that visual and auditory stimuli travel at different speeds, as well as the differences in processing and conduction times to the central nervous system. Stimuli that start at the same time do not necessarily produce the strongest interaction and varying the interval between stimuli can sometimes change enhancement to depression. In general, there is an inverse relationship between the strength of the stimuli and the response, with the combination of weak unimodal stimuli giving the strongest response when perceived together. In particular, the combination of two stimuli which are each imperceptible on their own can often be perceived.

Sensory stimuli presented to different sensory modalities are treated as a single multisensory object if they are perceived as synchronous, but not otherwise. Judgements of whether or not auditory and visual stimuli are simultaneous depend largely on their relative location, with stimuli from the same position more likely to be considered simultaneous than those from different positions and a large temporal window for perception of simultaneity (Zampini *et al.* 2005). Therefore judgements about the temporal order of stimuli are more accurate when they are spatially separated than when they are in the same position. However, differences in speed and temporal processing mean that auditory and visual signals from the same object are not received at the same time away from the horizon of simultaneity and therefore have to be combined into a single multisensory object resulting in the loss of the information on the relative temporal onsets of the different stimuli (Spence *et al.* 2003).

2.1.4 Multisensory Perception in the Superior Colliculus

Convergence of visual, auditory and somatosensory (tactile and relating to the internal organs) stimuli occurs at several places in the central nervous system, including the superior colliculus, which is discussed in more detail in Section 2.3.6. Visual, somatosensory and auditory representations in the superior colliculus have similar map-like representations and therefore eye movements due to non-visual stimuli are directed to the source of the stimuli, as in the case of visually evoked eye movements. However, they are generally less accurate, as auditory receptive fields are larger than visual ones.

Activation of a particular region of the superior colliculus produces movements of the eyes, ears and head to focus on the area in space represented by the stimulation site. Therefore, the motor circuitry of the superior colliculus has a map-like organisation. The superior colliculus has a number of different layers. The organisation of both visual and somatopic stimuli is map-like with a spatial relationship between the location of the stimulus and a point on the superior colliculus. There is a very regular relationship between the visual and somatotopic maps in the deep layers, with the visual map emphasising the central retina, which has the greatest visual sensitivity, and the somatosensory map emphasising the face and hands, which have the greatest tactile sensitivity. There is also an auditory map, but it is tonotopic, that is, based on sound frequencies or tones, rather than the location of the stimulus, as with the visual and somatotopic maps. Therefore, the spatial auditory map uses differences between the intensity and timing of sounds at the two ears. This is because visual and somatosensory peripheral nerve fibres respond to a stimulus in a restricted area, regardless of its intensity, whereas auditory nerve fibres respond to loud sounds regardless of their spatial location. The spatial auditory map is oriented similarly to the visual and somatosensory maps.

The visual, auditory and somatosensory representations in the superior colliculus use common axes and the auditory and somatosensory components of the multisensory representation can be described in visual coordinates and *vice versa*. Other than a general increase in receptive field size with the number of sensory modalities, there are few differences between unimodal, bimodal and trimodal neurons. Therefore their large receptive fields mean that the receptive fields of a visual-auditory neuron are rarely completely coincident and the three receptive fields of a trimodal neuron are almost never coincident. This lack of precision may mean that slight shifts in the peripheral sensory organs have little effect on the overall representation. The different sensory maps in the superior colliculus use many of the same neurons, so they may comprise one integrated multisensory map, rather than three parallel, but independent maps.

2.1.5 Studies of Multisensory Perception

There have been a number of studies of multisensory perception and the results of a few of them are summarised briefly here. There has been found to be strong crossmodal integration in motion perception, which sometimes produces an apparent reversal of the true direction of auditory motion, although this direction is clear when the auditory signal is presented on its own (Soto-Faraco *et al.* 2002). A sound and light are often perceived to be moving in the same direction even when they are in reality moving in opposite directions (Soto-Faraco *et al.* 2002; Zapparoli and Reatto 1969). In the case of auditory and visual motion, it is generally visual motion that captures auditory motion. However, in the case of simultaneous auditory and tactile motion, either modality is able to capture the other, though tactile motion has a stronger impact on auditory motion perception than *vice versa* (Soto-Faraco *et al.* 2004a). Temporal ventriloquism also occurs, with the presence of sounds affecting visual perception in the temporal domain (Morein-Zamir *et al.* 2003). Audiovisual interactions can induce perception of visual stimuli in the impaired hemifield of people with hemianopia (blindness in half the visual field). Associating a sound to a visual stimulus has been found to improve the ability to consciously perceive visual stimuli in the impaired visual field, but only when the sound and visual signal are spatially coincident (Frassinetti *et al.* 2005).

2.2 The Visual System

2.2.1 Introduction

The eye is a slightly asymmetrical sphere set in a protective cone-shaped cavity in the skull called the orbit or socket. It has a lengthwise diameter of about 24 mm and a volume of about 6.5 cc. The orbit is surrounded by layers of soft fatty tissue, which protect the eye and enable it to turn easily. A cross sectional view through the eye shows the following three different layers (see Figure 2.1):

1. The external layer or fibrous tunic; formed by the sclera and the cornea. The fibrous tunic provides the eye's shape when inflated by the pressure of the internal fluid (the intraocular pressure).
2. The intermediate layer; consisting of the iris and ciliary body at the front and the pigment epithelium and the choroids at the back.
3. The innermost layer; consisting of the retina, which is the sensory part of the eye.

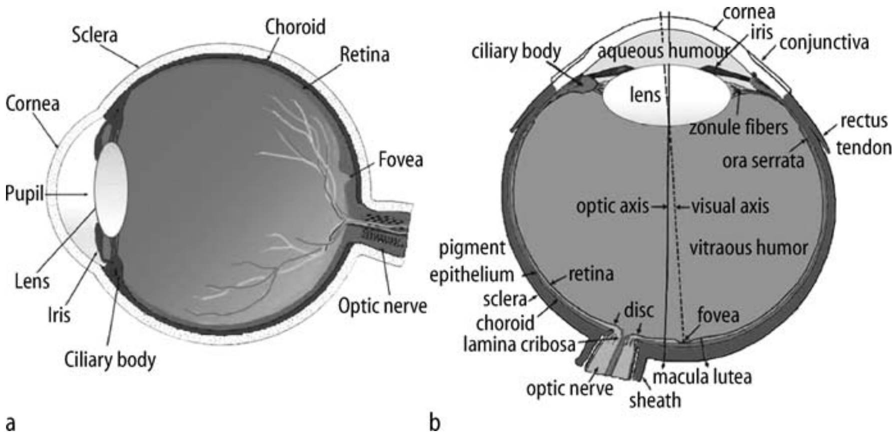


Figure 2.1a,b. Sectional views of the eye: **a** sagittal vertical section of the adult human eye; **b** sagittal horizontal section of the adult human eye. (Images reproduced by kind permission of Dr. Helga Kolb, from http://webvision.med.utah.edu)

The eyeball is divided into three fluid filled compartments:

1. The anterior (front) chamber (between the cornea and the iris), containing aqueous humour.
2. The posterior (back) chamber (between the iris, the zonule fibres and the lens), which contains the aqueous humour. The anterior and posterior chambers together comprise the anterior compartment.
3. The posterior compartment, containing the vitreous chamber (between the lens and the retina) which is filled with vitreous humour. The posterior compartment is much larger than the anterior compartment.

The optics of the eye have the function of focusing a clear image onto the retina. Light from an object enters the eye through the cornea, which is a clear dome covering the front of the eye. Together with the crystalline lens, it refracts or bends the light rays to focus them on the retina. Two-thirds of the refractive power or bending of the light rays is performed by the cornea and the remaining one-third by the lens. Although most of the refractive power of the eye is due to the cornea, the lens is responsible for the fine adjustments that allow the eye to focus on objects. This is because the lens, but not the cornea is able to change its shape. In the process of accommodation, the ciliary muscles at the side of the lens contract to make it more spherical to focus on near objects and relax to make the lens flatter when focussing on distant objects. The amount of light entering the eye is controlled by two sets of muscles, which adjust the size of the pupil. The tear glands produce fluid that keeps the surface of the cornea clean and lubricates it. The production of tears is stimulated by the regular blinking or closing of the eyelids. This blinking occurs at a greater rate when stimulated by irritation due to, for instance, a dust speck on the cornea.

Six extraocular muscles regulate the movements of each of the two eyeballs to allow them to track and fixate on objects. These muscles are arranged in three pairs on each eyeball, with the muscles in each pair working in opposition, and each pair of muscles controlling movements in one of three perpendicular planes. The two eyes need to track objects together with a precision of a few minutes of arc to avoid double vision.

The retina is the sensory part of the eye. It is about 0.25 mm thick, plate shaped and consists of ten layers, comprising three layers of nerve cell bodies separated by layers containing synapses or connections between the nerve cells, as well as inner and outer linings and the pigment epithelium. The photoreceptors are in the back layer of the retina, so that light has to pass through the other layers to stimulate them. There are two types of photoreceptors, rods and cones. The cones are responsible for daylight or photopic vision and colour vision, whereas the rods are responsible for night or scotopic vision, as they are much more sensitive to low intensities of light than the cones. There are three different types of cones, with each type of cone most sensitive to a particular frequency of light. This gives rise to colour vision, with the relative stimulation of the three cones determining the particular colour. There is only one type of rod and consequently people are not able to see colours at night in the absence of artificial light.

The rods and cones in the retina change the light energy into electrical and then chemical signals which are transmitted to the cells in the inner nuclear and retinal ganglion layers and thence to the optic nerve. The optic nerve connects the retina to the lateral geniculate nucleus (named after its shaped like a bent knee) from where the optic nerve fibres fan out in a broad band called the optic radiation into the interior of the brain to end in the primary visual cortex or striate cortex. Between the retina and the lateral geniculate nucleus the optic nerve fibres divide and half of them cross over at the optic chiasm. This separates the right and left visual worlds, so that the output from the right side of the two retinas goes to the right lateral geniculate nucleus and hence the right hemisphere of the brain and output from the left side of the two retinas goes to the left hemisphere of the brain. However, the image on the retina is reversed by the lens and therefore the right visual field, namely everything to the right of a vertical line through the point being looked at, is connected to the left hemisphere of the brain and the left visual field to the right hemisphere. Therefore, a massive stroke in the left side of the brain would lead to blindness in the right visual field of both eyes.

There are a number of other visual areas in the brain, which receive signals from the primary visual cortex and/or the lateral geniculate nucleus. Although the details of how processing in the brain takes place are not yet (fully) understood, it is known that different types of visual processing, such as of colour and movement, take place in different parts of the brain.

2.2.2 The Lens

The compound lens that projects an optical image onto the retina comprises the crystalline lens (which will subsequently be referred to solely as the lens) and the cornea. The anterior surface of the lens is covered by aqueous humour and its posterior surface is in contact with the vitreous body. Both these junctions have very low changes in refractive index, giving the lens only one-half the refracting power of the cornea despite its greater curvature and thickness. The relatively higher refractive power of the cornea is due to its curvature and high refractive index compared to air. However, it is the lens and not the cornea that provides the focusing power of the eye.

The lens is biconvex, transparent and semisolid. It has a slightly yellowish tint that increases with age. It has a diameter of 9–10 mm and a thickness of 4–5 mm, but both these dimensions vary with the focus of the eye on near or far objects. The lens consists of elongated cells called lens fibres and is enclosed in an elastic membrane called the lens capsule. New fibres are laid down on top of older ones, which become denser and more tightly packed. The nucleus of the lens consists of the older fibres at the centre of the lens and the lens cortex comprises the new fibres near the surface. Consequently, the lens has a hard nucleus surrounded by a soft cortex. The lens does not shed any cells and therefore it continues to grow as new cells are laid down, but the rate of growth decreases with age. The chemical composition of the lens is not uniform, but on average it is 63% water, 36% protein and 1% lipids, salts and carbohydrates.

Changes in the thickness and curvature of the lens are achieved by relaxation and contraction of the ciliary muscle which affects the tension of the suspensory ligament called the zonule of Zinn, which is attached between the lens and the ciliary muscle. Relaxation of the ciliary muscle puts the zonule under tension and pulls on the lens capsule, making it thinner to focus the eye on distant objects. Contraction of the ciliary muscle relaxes the zonule and release the tension of the lens capsule, so that it rounds and thickens due to its natural elasticity, focusing the eye for nearer objects. As people become older, the lens becomes more curved for a given amount of accommodation to compensate for its decrease in refractive index and there may be internal compensatory changes in refraction. The angle of attachment of the zonules changes with the size of the lens. Eventually the force from the zonules becomes approximately tangential to the surface of the lens, so that its relaxation has less and less impact on the lens shape and the lens system consequently gradually loses its ability to accommodate to focus on objects at different distances. This is called *presbyopia*.

The lens cells consist of protein molecules in a colloidal solution. The low concentration and small size of the protein molecules result in little scattering of light, so that the majority of the incident light is transmitted, making the lens transparent. A loss of transparency of the lens, particularly if it reduces vision, is called a cataract. The lens does not have a blood supply and is provided with nutrients by the aqueous humour.

2.2.3 The Iris and Pupil

The iris and pupil regulate the amount of light entering the eye and illuminating the retinal photoreceptors. The pupil is slightly below and to the nasal side of the centre of iris. The diameter of the pupil is controlled by the iris muscles, the sphincter iridis, and the dilator pupillae. It can vary from less than 1 mm to more than 9 mm. To achieve this considerable change in size the iris is able to reduce to about 13% of its maximum length, a much greater degree of contractibility than any other muscles. The iris is constantly in motion, even when there are no changes in illumination and accommodation (focusing of the lens). Its movements are largest in moderately bright light and the frequency of oscillation increases with the light intensity. The resulting instability of the pupil is most easily seen in young people. About one-fifth of the population have small differences of about 0.4 mm in the size of the pupils of the two eyes without any associated visual impairment or disease of the eye.

Stimulation of one eye with light leads to a contraction of the pupil in the other eye due to the division of the optic nerve fibres at the optic chiasm, so that the axons from both the left and right half fields are connected to the right and left hemispheres, respectively. The pupil also takes part in the near reflex, which involves the three functions of convergence of the two eyes, accommodation of each eye and constriction of both pupils when viewing a near object, even when the illumination does not change. As well as reducing retinal illumination, pupil constriction reduces blurring of the image. The near reflex results from the sphincter pupillae, ciliary and medial recti muscles all working together to improve the image, its focus on the retina and the depth of field.

2.2.4 Intraocular Pressure

The eye is an inflated approximately spherical shell. Since the cornea and sclera are thin connective tissues with little rigidity, the shape and optical functions of the eye are maintained by the intraocular pressure or pressure in the eye of about 15 mm Hg. This pressure is determined by the amount of aqueous humour in the eye and counterbalanced by the tension in the outer tunic of the eye. The main source of the aqueous humour is the blood flowing through the ciliary arterial system. The processes by which the clear transparent colourless aqueous humour is produced from the blood are not fully understood and suggestions include filtration (to remove cellular components), dialysis (to separate smaller from larger particles in solution by diffusion through a membrane) and secretion (with cells using metabolic energy to modify material). The rates of production and drainage of aqueous humour need to be the same to maintain the constant size of the eye.

The aqueous humour drains from the eye into the veins *via* a circular channel, called the canal of Schlemm through the trabecular network of connective tissue. The intraocular pressure is determined by the resistance to flow of these tissues combined with the rate of production of the aqueous humour. This is a channel in the corneal stroma around the eye, which is located at the filtration angle between the cornea and the eye. Blockage of the filtration angle and consequent closure of Schlemm's canal, reduces the outflow rate of the aqueous humour, resulting in a rapid and painful increase in intraocular pressure, which often leads to angle-closure glaucoma. A slow increase in the production rate of aqueous humour or the liquid flow resistance in the outflow pathway will lead to a slow rise in intraocular pressure and possibly also to open-angle glaucoma.

Temporary increases in intraocular pressure may result from stress or anxiety, wearing a tight collar or holding one's breath. There is a variation of about 4 mm Hg from the early morning high to the evening low and a variation of about 2 mm Hg from 15 to 17 mm Hg with age. The aqueous humour has a greater concentration of dissolved substances than the blood, leading to an osmotic flow of water across the blood aqueous barrier. Medicaments can be used in acute-angle glaucoma to manipulate this flow rate to lower the intraocular pressure quickly to avoid damage to the eye.

2.2.5 Extraocular Muscles

Description of the movement of the eyes can most easily be done using a frame of reference that defines the axes of rotation of the eyeball or globe. Its centre of rotation depends on the position of the eyeball in the orbit. Torsional movements take place about the visual axis. In intorsion and extorsion the upper part of the eye rotates towards and away from the nose. Horizontal rotations take place about a vertical axis and vertical rotations about a horizontal axis. Elevation and depression involve upward and downward movements of the eye, whereas adduction and abduction involve horizontal rotations towards and away from the nose.

There are three pairs of extraocular muscles (see Figure 2.2):

1. The inferior and superior rectus, whose main actions are elevation and depression of the eye, with secondary actions of adduction and torsion.
2. The superior and inferior oblique, whose main actions are intorsion and extorsion. Their secondary actions involve horizontal and vertical displacements with increasing adduction.
3. The lateral and medial rectus, whose main actions are adduction and abduction, namely, horizontal rotation towards or away from the nose. They have minimal secondary actions.

Controlling eye movements requires sensory feedback about the current eye position and the ability to move the eyes either quickly or slowly, as well as to keep them relatively stationary. The extraocular muscles are controlled by three pairs of cranial nerves:

1. The oculomotor nerve (cranial nerve III)
2. The trochlear nerve (cranial nerve IV)
3. The abducens nerve (cranial nerve VI)

The oculomotor range is about 90°, *i.e.* a quarter of a circle; this is the angle through which the direction of the gaze can be moved by moving the eyes without moving the head. When the visual axes are directed straight ahead, the eyes are in the primary position. Secondary and tertiary positions are obtained by purely horizontal and vertical and combined horizontal and vertical movements respectively from the primary position. Strabismus or squint occurs due to misalignment when the observer tries to fixate an object with both eyes.

There are a number of different ways of classifying eye movements. Conjugate eye movements displace the visual axes of both eyes in the same direction as though they were yoked together, whereas vergence movements move the visual axes in opposite directions, either towards (convergence) or away from (divergence) each other. Conjugate movements can be vertical, horizontal or oblique. Another important distinction is the speed, with the two main classes of quick movements, such as smooth pursuit, and slow movements, such as saccades. The eyes are only able to detect motion or relative motion and they are therefore constantly making small tremors, small drifts and periodic microsaccades, called physiological nystagmus to enable stationary objects to be seen. This requirement for relative movement for vision has the benefit of preventing the shadows cast by the blood vessels obscuring the object's image on the retina.

Slow eye movements are used to stabilise the retinal image, though not to the extent that it disappears. Smooth pursuit movements are used to track the movement of an object, such as a train or an animal, and generally cannot be performed voluntarily when there is no moving target. These movements stabilise the image of the moving object, while the image of the background keeps moving on the retina. The smooth pursuit system uses a predictive strategy, which generates an internal replica of the target velocity and matches eye velocity to it over a period of time.

Eye movements have very fine tolerances to ensure proper alignment and coordinated movements of the two eyes, giving a need for precise control of the extraocular muscles. This involves subsystems that are under voluntary control, but which do not require conscious thought. Conjugate movements of the eyes generally involve relaxation or contraction of the opponent muscles of a pair in both eyes. For instance, a conjugate movement to the left requires simultaneous contraction of the lateral rectus muscle of the left eye and the medial rectus muscle of the right eye.

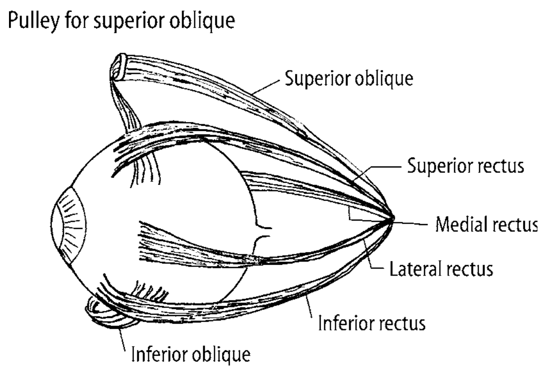


Figure 2.2. Extraocular muscle pairs

The movement of objects or the eyes (other than the small continuous movements mentioned above) leads to displacement of the image on the retina. The image is stabilised by reflex and voluntary movements which counteract the image motion. For instance the vestibular-ocular reflex opposes image displacements due to head movements to give a smooth following movement in the same direction called an optokinetic movement. Large optokinetic movements are generally followed by a quick, anti-compensatory movement of the eye toward the middle of the oculomotor range after which the slow movement starts again. This combination of slow and fast movements is a form of nystagmus or to-and-fro movement of the eyes.

Quick eye movements, such as voluntary and reflex saccades, are used to move the retinal image onto the foveas, for fine detail vision. They are a response to the appearance of an object of interest away from the current point of fixation. The oculomotor system creates saccadic eye movements to correct the retinal error or separation of the image from the fovea in both distance and direction. Current models of saccade control are based on tracking control, though the mechanisms are not fully understood.

2.2.6 Eyelids and Tears

The tear film of the eye provides a very smooth optical surface on the cornea, lubricates the eye and eyelids, keeps the cornea moist and protects the eye from infection by destroying bacteria. It consists of an outer oily layer, a central aqueous layer, which is the main component, and a layer of mucin (large molecules consisting of both protein and carbohydrate). The aqueous part of the tear film continuously evaporates and is replaced, so that the concentration of the tears varies considerably. People blink about once every 5 s. When the eyelids move down over the eye, they wipe away some of the tear film, reducing the thickness of the layer. This fluid is then spread uniformly across the eye when the lids open.

The lids protect the sensitive eye tissue and a reflex blinking reaction closes the lids in response to irritation or possible danger to the eye. The lids also spread the tears uniformly over the eye to maintain a high quality optical surface on the cornea and they close to reduce exposure to light during sleep. Strong light, loud noises, touching the cornea and objects approaching the eye unexpectedly can stimulate involuntary blinking. There is some research evidence that the main blink characteristics are affected by hard contact lenses, whereas soft lenses do not have any effect (Fatt and Weissman 1992).

2.3 Visual Processing in the Retina, Laternal Geniculate Nucleus and the Brain

2.3.1 Nerve Cells

The brain and retina consist mainly of nerve cells or neurons. Like other cells, neurons have a globular cell body, the perikaryon or soma, and contain a nucleus, mitochondria and other organelles (subcellular structures with a specialised function). Signals are transmitted between neurons by nerve fibres or axons which are shaped like long thin cylinders and connected to the cell body. Dendrites are branching, tapering fibres that leave the cell body, generally in a different direction from the axon. The cell body, axon and dendrites are enclosed in a cell membrane. Axons range from less than 1 mm to 1 m in length, whereas dendrites are generally in the millimetre range. Axons usually split near their ends into many branches that form synapses or connections without quite touching the cell bodies or dendrites of other neurons. Signals in a nerve are initiated close to where the axon joins the cell body and travel along the axon toward the synapse with the next neuron. At the synapse, chemical transmission is used to transfer the signal from the presynaptic cell to the postsynaptic cell. There are a number of different types of neurons, with between one hundred and one thousand different types in the brain.

Almost all the cells in the nervous system receive inputs from more than one cell, called convergence, and supply outputs to possibly hundreds or thousands of other neurons, called divergence. The term grey matter is used for nerve cell bodies, their dendrites and blood vessels, which have a greyish brown colour, whereas the term white matter is used for axons covered with a myelin sheath, which is a white, fatty insulating material. The cerebral cortex consists of grey matter, whereas the inner portion of the cerebrum comprises white matter.

The term receptive field is used to describe the region of space in which the presence of a stimulus will affect a neuron. For instance, the receptive field of a photoreceptor is a cone shaped volume of space covering all the visual directions in which light will affect the firing of the cell. The receptive field is often given as the region of the retina where the action of light affects the firing of the neuron. Visual neural cells can be classified in terms of the properties of their receptive fields as on-centre, off-centre and on-centre, off-surround. On-centre cells respond to the onset of light and are inhibited by its offset and therefore respond to small spots of light. Off-centre cells are inhibited by the onset of light and excited by its termination. On-centre off-surround cells respond to both light onset and offset, that is to a change in the light level, whether it is an increase or a decrease.

On-centre cells respond at a significantly increased rate when a small spot of light stimulates the centre of the receptive field. The response increases until the centre of the receptive field is filled and then decreases as the surround is included, since stimulating the surround has an inhibitory effect. Off-centre cells have their greatest response to a black spot on a white background, as only the surround and not the centre is being stimulated. The eye responds to brightness differences between objects and their surrounds, rather than absolute brightness and this is expressed by centre-surround receptive fields. Centre/surround or lateral antagonism is used to compare light levels. This is a property of both bipolar and ganglion cells, but is stronger in ganglion cells. The existence of both on- and off-centre cells allows people to see both dark and light objects. The centre/surround organisation allows the response of part of the cell to be modified by stimulation of a neighbouring area.

2.3.2 The Retina

The retina consists of nervous tissue. It converts light energy to chemical energy and then to electrical energy to transmit visual signals to the brain for processing. The retina is part of the brain that separated from it early in development, but maintains a connection through the optic nerve. It has the shape of a plate and varies in thickness from about 0.1 to 0.5 mm. It is thickest round the optic disk or nerve head, which is the area at the back of the inside of the eye where the optic nerve connects to the retina. It thins at the *ora serrata*, the equator and fovea. The central portion, which is about 5–6 mm in diameter, contains the *macula lutea* (yellow area) and fovea, which is the site of most acute vision. The peripheral retina extends back from the central retina to the *ora serrata*, which is the serrated end of the optic part of the retina that forms the junction between the retina and the ciliary body. It is located a little behind the ciliary body and marks the limits of the sensitive portion of the retina.

The retina is divided into 10 layers and these can be classified as follows (see Figure 2.3):

- The (retinal) pigment epithelium, which is the outermost layer.
- Three main layers of cells, the outer nuclear, inner nuclear and ganglion cell layers, consisting of the photoreceptors; bipolar, horizontal and amachrine cells; and ganglion cells respectively.
- The outer and inner plexiform layers, containing the synapses between the axons and dendrites of the cells of the three main layers.
- Inner and outer limiting membranes, though it should be noted that the outer limiting membrane is the third layer of the retina.
- The optic nerve fibre layer; containing the axons of the ganglion cells.

The photoreceptors are at the back of the retina and therefore incoming light has to pass through the other layers before reaching the photoreceptors. A row of cells, the pigment epithelium, containing the black pigment, melanin, also found in the skin, is located behind the photoreceptors. This melanin absorbs light that has passed through the retina to prevent it from being reflected back and scattering around inside the eye. The melanin containing cells also help to restore the light-sensitive visual pigment in the photoreceptors after they have been bleached by light. Both these functions require the melanin pigment to be close to the photoreceptors and may be the reason for the location of the photoreceptors at the back rather than the front of the retina.

There are two types of photoreceptors, rods and cones. There are about 100–110 million rods and 6–8 million cones. The rods are responsible for scotopic vision in dim light and the cones for photoptic vision in bright light. The distribution of rods and cones varies considerably over the retina. The cones have their greatest density at the fovea and their density decreases rapidly towards the periphery. There are no rods at the fovea and the density of the rods increases away from the fovea to a maximum at 20° eccentricity and then decreases again towards the periphery of the retina.

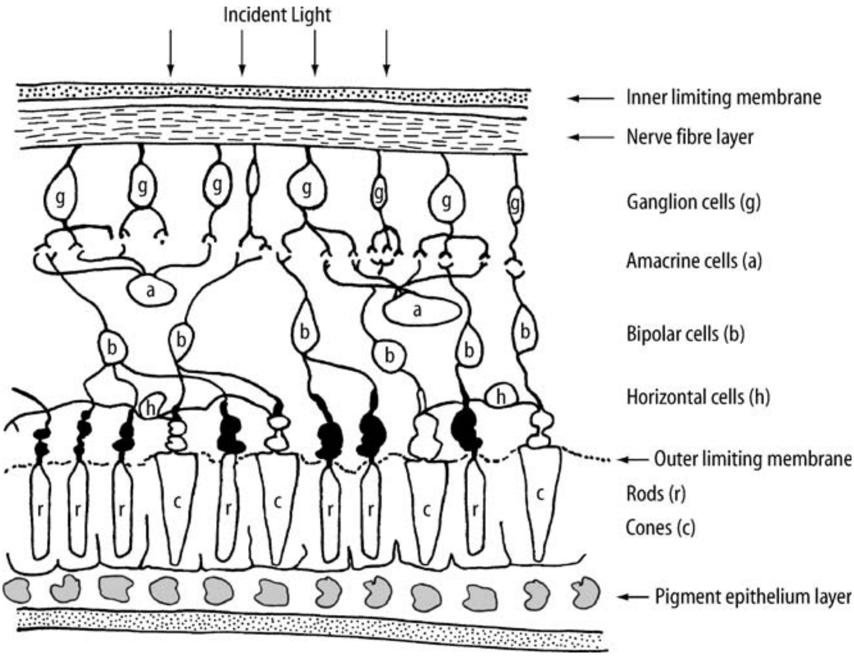


Figure 2.3. Sketch schematic of a section through the retina

The fovea at the centre of the retina has only cones and the best fine detail vision. To avoid any possible blurring of the image from the other layers of the retina (though they are fairly transparent), these layers are displaced to the side at the fovea to give a ring of thicker retina, with the central cones exposed in the shallow pit of the fovea at the very front.

Rods and cones both have an outer segment, containing the visual pigment and a stack of about 1000 horizontal circular discs, an inner segment, a cell body region, containing the cell nucleus, and a terminal. Cones frequently have a conically shaped outer segment and rods a cylindrically shaped outer segment. However, outer segment shape is not an infallible guide to photoreceptor type since, in the fovea, the cone outer segments are more slender and longer than those of the rods in order to increase their packing. The definitive distinguishing feature is the fact that all the cone discs remain attached to the outer segment membrane while those of the rods are freely floating, apart from the few at the very base of the outer segment. The other distinguishing feature is that rods have small axonal terminals called spherules while cones have greatly expanded foot-like terminals called pedicles. However, the most significant difference between the rods and cones is that rods are sensitive to very dim light, whereas cones require much brighter light. Rods and cones contain light-sensitive pigments, with all rods having the same pigment and the three different types of cones each containing a different pigment. These four pigments are sensitive to different wavelengths of light.

Information from the rods and cones has both a direct path *via* the bipolar cells to the approximately one million ganglion cells and an indirect path that involves the horizontal and amacrine cells. The horizontal cells use long connections parallel to the retinal layers to link photoreceptors to bipolar cells. The axons from the retinal ganglion cells pass across the front of the retina in the nerve fibre layer, come together at the optic disc and leave the eye to form the optic nerve. In the direct path, each bipolar cell has input from only one or a small number of photoreceptors and each ganglion cell has input from only one or a small number of bipolar cells, whereas the indirect path is extended through wider lateral connections. In the fovea, where fine detail vision is best, a single cone feeds *via* one bipolar cell into a single ganglion cell, whereas the convergence of photoreceptors onto bipolar cells and ganglion cells increases from the fovea towards the periphery of the retina.

Retinal ganglion cells can be divided into small P cells (from the Latin *parvus* for small) and large M cells (from the Latin *magnus* for large) and also into cells with on- and off-centre responses. The P cells predominate in the fovea. They have small receptive fields and are responsible for fine detail vision. The M cells have large dendritic arrays and larger receptive fields and are mainly found in the peripheral retina. They are sensitive to motion, but insensitive to detail and colour. This separation of visual processing into colour, form and detail; and motion and spatial processing is maintained throughout the visual system and is discussed in more detail in Section 2.3.7.

2.3.3 The Optic Nerve, Optic Tract and Optic Radiation

The optic nerve transmits visual information in the form of electrical impulses from the retina to the lateral geniculate nuclei in the brain. Although it is the 2nd of the 12 paired cranial nerves, it is considered to be part of the central nervous system. The axons of the retinal ganglion cells are the main constituent of the optic nerve, but it also contains glial support cells and nerve tissue.

Glial cells are non-neuronal cells that provide maintenance, support and nutrition in the central nervous system. They comprise about half the weight of the brain and are about 10 times as numerous as the nerve cells they look after. They try to keep the neurons healthy by producing new myelin when it is damaged and laying down scar tissue. They remove dead cells and other debris from the central nervous system and chemically regulate the levels of neurotransmitter or chemical signals between different neurons.

The optic nerve starts at the optic nervehead or optic disc, which is connected to the retina close to the macula. There is a blind spot at the optic disc, due to the absence of retinal sensory receptor cells, but this is rarely noticeable due to overlapping vision from the two eyes. The optic nerve fibres leave the retina through pores in the lamina cribrosa, which is a sieve-like structure in the optic nerve head. It is also the point at which blood vessels enter and leave the eye. The nerve fibres form a rim round the optic nerve head and are absent from the central indentation called the optic cup. The blood supply to the optic nerve is mainly from blood vessels originating in the ophthalmic artery.

The nerve fibres in the retinal nerve fibre layer and the optic nerve head are unmyelinated, to ensure that the retina is sufficiently transparent for light to reach the photoreceptors in the outer retina, but the rest of the optic nerve is covered in a myelin sheath which insulates the axons. In myelinated nerves impulses are conducted by saltatory (jumping) conduction between the nodes of Ranvier, where depolarisation occurs. This makes signal transmission much faster than in unmyelinated nerve fibres, where conduction of impulses is carried by depolarisation, and also conserves metabolic energy. Since the optic nerve is part of the central rather than the peripheral nervous system, its fibres cannot regenerate if they are damaged. Therefore, damage to the optic nerve results in irreversible blindness.

At the optic chiasm there is a partial decussation (crossing) of fibres from the temporal (side) visual fields of both eyes, so that information from the right visual hemifields of both eyes goes to the right lateral geniculate nucleus and from the left visual hemifields of both eyes to the left lateral geniculate nucleus (see Figure 2.4) After the optic chiasm, the optic nerve fibres become the optic tract. All the axons of the optic tract synapse or make connections in the lateral geniculate nucleus.

From the lateral geniculate nucleus axons travel through the deep white matter of the brain to the visual cortex in the occipital lobe as the geniculo-calcerine tract or optic radiation. Most of the fibres travel directly to the visual cortex, but fibres carrying information from the lower retina loop round the lateral ventricle (a curved cavity in the front part of the forebrain) as Meyer's loop and pass into the temporal lobe (part of the cerebrum which is in front of the occipital lobe). Therefore a lesion in the temporal lobe can cause blindness in the upper outer quadrant of vision, whereas a lesion in the optic radiation going directly to the

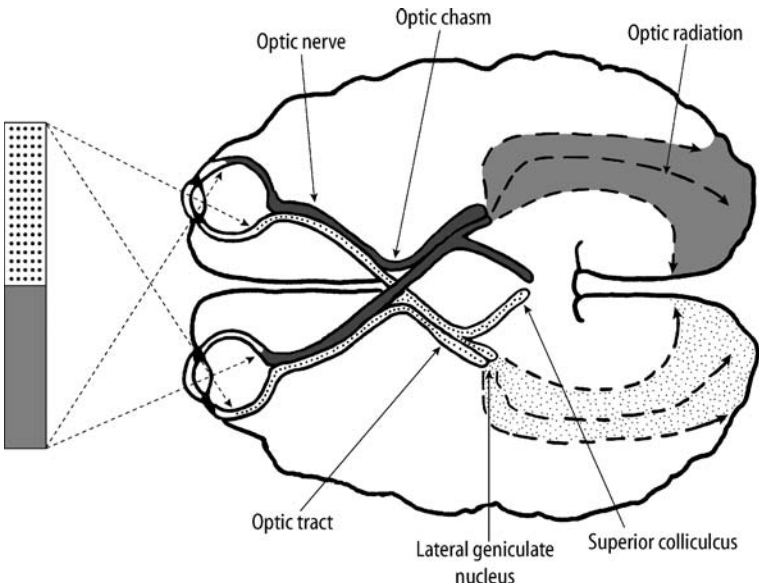


Figure 2.4. Visual pathway, showing the partial crossing of optic nerve fibres at the optic chiasm

visual cortex produces homymous hemianopia or blindness in the contralateral visual field of both eyes.

2.3.4 The Lateral Geniculate Body or Nucleus

There are two lateral geniculate nuclei (LGN), one on the right and one on the left side, and they are subdivisions of the cerebral hemispheres. They are folded, giving them the appearance of bent knees and the name geniculate is derived from the Latin for knee. Each of the two lateral geniculate bodies consists of six layers, as is generally the case for the cortex. Both LGNs receive signals from retinal ganglion cells in both eyes. However the layers are organised so that layers 5, 3 and 2 receive input from optic nerve fibres from the same side (ipsalateral) eye, whereas optic nerve fibres connect layers 6, 4 and 1 to the opposite side (contralateral) eye. In addition, there is a detailed correspondence between the retina and the LGN, with adjacent points in the retina projecting to adjacent points in each layer of the LGN.

The six layers are organised into two groups. The outer (dorsal) group of four layers has small cell bodies, which are called the parvocellular or P layers (from the Latin *parvus* for small). They receive input from the small ganglion cells of the retina. The inner (ventral) group of two layers has large cells bodies and is called the magnocellular or M layers (from the Latin *magnus* for large). The P cells are concerned with colour vision, whereas the M cells are not. Like the retinal ganglion cells, the lateral geniculate cells have receptive fields with a centre-surround organisation and can be categorised according to whether their receptive fields have on-centres or off-centres and their detailed colour properties.

2.3.5 The Primary Visual or Striate Cortex

The primary visual cortex is located in the occipital lobe at the back of the brain. The occipital lobe is the visual processing centre of the brain and the smallest of the four brain lobes (parietal, frontal, temporal and occipital) which together with the insula compose the cerebral cortex. The primary visual cortex consists of a 2 mm thick plate with a surface area measured in centimetres and occupies most of the surface of the occipital lobe. It is organised into orientation columns perpendicular to the surface of the cortex, with all the cells in the column having the same or very similar preferred stimulus orientations. There are also ocular dominance columns in which all the cells are dominated by the same eye. The primary visual cortex has about 200 million cells, whereas the geniculate has about 1.5 million. It was the first area of the cerebral cortex to be distinguished, due to the distinct white stripe in cross section, which gives it the name of striate cortex. The striate cortex is Brodmann area 17. This numbering system is based on a somewhat arbitrary division of the cerebral cortex into about 50 different regions, with numbers given according to the order in which the German neurologist Brodmann studied them. The fact that this numbering system is still in use indicates a lack of understanding of the brain, which has prevented a better system being derived. The primary visual cortex is divided into a number of different layers, which differ in their inputs, local interconnections and the structures to which their outputs are connected.

The LGN cells project in a point to point manner to the primary visual cortex on the same side of the brain *via* the optic radiation. Therefore, the cortex receives input from the contralateral half of the field of view and adjacent retinal points are mapped into adjacent points on the cortex. This gives a topographic map, with each part of the retina represented in the primary visual cortex. The right visual cortex registers information in the opposite hemifield, *i.e.* the left hemifield of view, whereas the left visual cortex records the right hemifield of view. Similarly, the upper retina registers events in the lower field of view and is represented in the upper calcarine cortex, whereas the lower retina, which registers events in the upper field of view, is represented in the lower lip of the calcarine cortex. The pole of the occipital lobe represents central vision.

Cortical neurons respond most strongly to elongated stimuli, such as lines and edges, but not to diffuse light stimuli. They exhibit orientation tuning, namely, they respond most strongly to stimuli with a particular orientation. Tuning curves can be drawn for different cells to show the decline in response away from the preferred orientation. The function of orientation tuning is not yet understood. Cortical neurons have larger receptive fields than ganglion or lateral geniculate nucleus cells. Their receptive fields are locked to retinal locations and move with eye movement, giving the retinotopic map on the cortex surface.

Visual information comes from the border and not the interior of a shape, thereby reducing the number of cells required for visual processing, though it is still very large. The importance of movement sensitive cells is probably because changes in the world play a greater role in survival than static conditions. Since the eye is spherical, retinal magnification is constant, with about 3.5° of visual field per millimetre of retina. However, the receptive field centres of the ganglion cells are much smaller in and near the fovea and grow in size with distance from the fovea. Therefore, a much larger number of ganglion cells corresponds to each millimetre of foveal than peripheral retina. The cortex is folded to accommodate the additional connections to the fovea, while maintaining a constant thickness and column widths. The cortical column is a group of neurons (nerve cells) in the cortex which are arranged on top of each other, that is, as a column, and which have nearly identical receptive fields (or regions of space where the presence of a stimulus will affect the firing of the neuron).

2.3.6 The Extrastriate Visual Cortex and the Superior Colliculus

The areas outside the visual cortex which respond to visual stimuli are collectively known as the extrastriate visual cortex. They are connected to the striate cortex and each other by direct and indirect projections. Output from the striate cortex goes mainly to the neighbouring band of cortex called visual area 2 or Brodmann area 18 and which almost completely surrounds it.

It is believed that a basic analysis of the input from the lateral geniculate nucleus is carried out in the striate cortex and the signal is then transmitted to the prestriate cortex for further analysis in the visual areas surrounding the striate cortex. These areas are classified either as visual areas 2, 3, 3A, 4 and 5 (V2, V3, V3A, V4 and

V5) or Brodmann areas 18 and 19. Areas V2 and V3 are divided into dorsal and ventral (upper and lower) halves that wrap round V1.

The superior colliculus is situated in the midbrain. It has a role in multisensory integration of perception, as was discussed in Section 2.1.4. It comprises seven alternating cellular and fibrous layers and is generally divided into two parts, consisting of the upper three and the four deeper layers. It is only the deeper layers that receive inputs from different sensory modalities, as well as from motor related structures. It is also the deeper layers that are most involved in using sensory information to affect motor behaviour. Visual organisation in both the superficial and deep layers of the superior colliculus is map-like or visuotopic, but the deep layer map is coarser than the superficial one and also includes the far periphery of the retina.

The functions of the superior colliculus include the control of eye movements and the direction of the gaze, using saccades and sometimes also head movements. In particular, it seems to be involved in the mechanisms that direct the gaze towards a target and ensure that the image is placed on the macula.

2.3.7 Visual Pathways

It is generally believed that there are two main visual pathways or information streams, which are responsible for fine detail, colour and form; and motion and spatial location processing. The fine detail pathway is referred to variously as the parietal, ventral, P or parvocellular pathway and sometimes also as the What pathway. It connects the smaller P ganglion cells in the retina to the parvocellular layers in the lateral geniculate nucleus, continues to the striate cortex and terminates in the parietal lobe after passing through the visual areas V2 and V4 of the extrastriate cortex. It is believed to be involved in form recognition, object representation and fine detail and colour perception. It has been suggested, but not confirmed, that visual area V4 is the colour-processing centre of the brain.

The motion pathway is referred to as the parietal, dorsal, M or magnocellular pathway and sometimes as the Where pathway. It connects the larger retinal ganglion cells *via* larger diameter axons to the ventral magnocellular layers of the lateral geniculate body and then continues to the primary visual cortex to terminate in the parietal lobe after passing the visual areas V2, V3 and V5 or MT (middle temporal) region of the extrastriate cortex. It is sensitive to changes in luminance at low light levels and motion. It is believed to be involved in spatial processing, including motion, representing object locations and controlling the eyes and arms, particularly when visual information is used to guide saccades or reaching movements. There is a reasonable body of evidence to suggest that the visual area MT is the motion-processing centre of the brain.

The terms M and P, or magnocellular and parvocellular pathway, are generally used for the first part of the pathway from the retinal ganglion cells to the primary visual cortex, whereas the terms ventral or parietal and dorsal or temporal stream or pathway are generally used for the part of the pathway from the primary visual cortex to the parietal or inferior temporal lobe.

This division of the functions of these two pathways is consistent with clinical observations that lesions in the parietal lobe affect perceptions of spatial relationships, whereas those in the temporal lobe affect memory for objects and faces. However, lesions in both areas may produce transient impairments, implying that visual functions are widely distributed or other areas are able to take over functions previously carried out by areas which have acquired lesions. Therefore, the model of separate streams of visual processing concerned with 'what' and 'where' is a simplification, and cortical processing is more complicated.

There is also another important perceptual pathway from the eye, called the collicular pathway, which includes the superior colliculus, but which will not be discussed further here.

2.4 Vision in Action

2.4.1 Image Formation

Points on real objects reflect diverging light rays with negative light vergence. On entering the eye, the light rays are reduced by the aperture of the pupil, and therefore the divergence decreases with pupil size. The amount of refraction required to focus light on the retina increases with the vergence of light. The light vergence can be stated, using Figure 2.5, in terms of the angle ϕ of light from an object at distance d passing through the pupil of diameter a, as

$$\tan\left(\phi/2\right) = a/2d$$

so that

$$\phi = 2\tan^{-1}\left(a/2d\right)$$

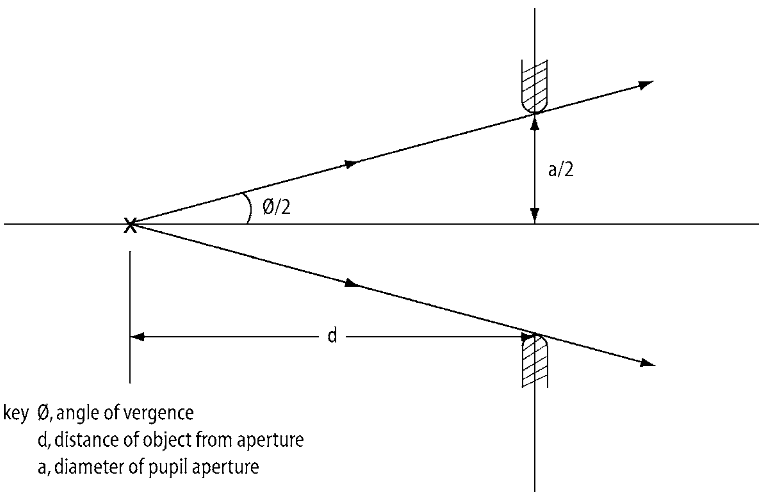


Figure 2.5. Vergence of light through the eye pupil

For a fixed distance, the vergence varies approximately linearly with the aperture size and approximately, inversely with the distance for a fixed aperture size.

Image formation by a lens results from refraction of light due to the fact that the speed of light varies with the medium it is in. Light moves more slowly in matter, such as the lens of the eye, than it does in a vacuum. The degree of reduction in speed depends on the type of material, with the ratio of the speed of light, c, in a vacuum to its speed in another medium, v, equal to the refractive index, n, of the medium. *i.e.*

$$n = c/v > 1$$

The speed of light in air is very close to that in a vacuum and therefore the refractive index of air is generally taken to be 1.0. The refractive index is also dependent on wavelength and greater for shorter wavelengths than longer ones. The change in speed with the change in medium generally also results in a change in direction. This change in direction leads to a bending of the light rays. The change in direction can be expressed in terms of Snell's law which states that the ratio of the sines of the directions (θ_i and θ_r for the incident direction and direction after refraction respectively) in two different media is equal to the ratios of the respective refractive indices (n_i and n_r):

$$\frac{\sin\theta_\mathrm{i}}{\sin\theta_\mathrm{r}} = \frac{n_\mathrm{i}}{n_\mathrm{r}}$$

Since the degree of bending or refraction is determined by the difference between the refractive indices of the two media, the air-cornea interface has the greatest effect on the eye's optical power. Vision is blurred underwater due to the reduced difference between the refractive indices of the cornea and water as compared to the cornea and air. Any scarring of the surface of the cornea or asymmetrical curvature can significantly reduce image quality, whereas the small loss in power due to removal of the lens can be compensated for by spectacles, contact lens or intraocular lens implants.

In general, the optical properties of the eye can be modelled by a lens of negligible thickness. Parallel rays striking the lens are made to converge towards a point on the axis called the focal point of the lens. A ray of light passing through the centre of a thin lens does not change direction. The nodal point is the point on the axis of the lens where rays of light pass through without changing direction. The eye is a compound optical system that has two nodal points. However, since the two nodal points are close together relative to the distance from the eye to the object being viewed, they can be treated in practice as one nodal point about 15–17 mm in front of the retina. This distance is called the posterior nodal distance and varies slightly between people. Using the one nodal point approximation makes it easier to calculate the size of the image on the retina. From Figure 2.6, the angle ϕ is given by

$$\phi = \tan^{-1}\left(\frac{y}{x+d}\right) \approx \tan^{-1}\left(\frac{y}{x}\right) = \frac{\text{height of object}}{\text{distance of object from eye}}$$

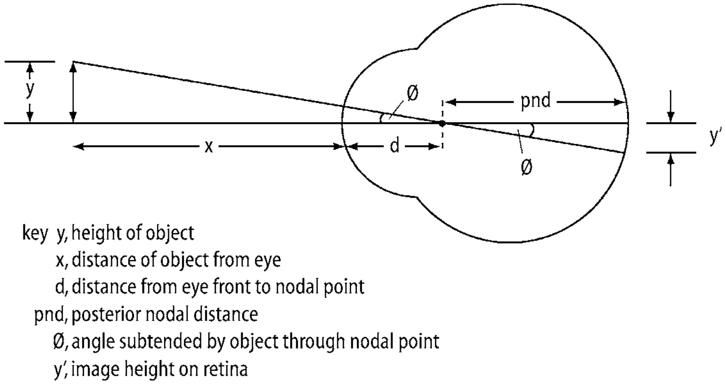


Figure 2.6. Simple model for eye optics using one nodal point

The height of the retinal image y' is given by

$$y' = \tan\phi \times \text{posterior nodal distance}$$
$$= \text{posterior nodal distance} \times \left(\frac{y}{x}\right) = 17 \times \left(\frac{y}{x}\right) \text{ mm}$$

2.4.2 Accommodation

Accommodation is the process of changing the curvature of the lens, generally at the front surface, to change the angle of refraction of impinging light. When the ciliary muscle moves closer to the lens, pressure on the zonules relaxes and near objects are focused on the retina (see Figure 2.7b). When the ciliary muscle is relaxed, tension on the zonules and the lens capsule increases and rays of light from distant objects are focused on the retina (see Figure 2.7a). The accommodation reflex seems to be controlled by a combination of negative feedback and a tonic controller to minimise blur of the retinal image. When this blur exceeds a threshold, a neural signal to the ciliary muscles stimulates or inhibits it to correct the focusing error. This produces clear vision quickly, but not in the long-term, as the neural signal decays rapidly. A slow or tonic controller provides sustained and self-adaptive regulation of accommodation by giving long-term shifts in the tone of the ciliary muscle. When there is no retinal image, for instance in the dark, the tonic controller generally adds about 1.5 D of accommodation to the lens, to make it sufficiently curved to give a sharp focus to an object at about 67 cm. The ability to adapt to distant targets may have a role in preventing work-induced myopia, as it leads to relaxation of the ciliary muscle which is useful after a long period of close work.

2.4.3 Response to Light

When exposed to light, the visual pigments absorb photons and are chemically changed into other compounds that absorb light less well or have a different wavelength sensitivity. This process is called bleaching, as the photoreceptor pigment

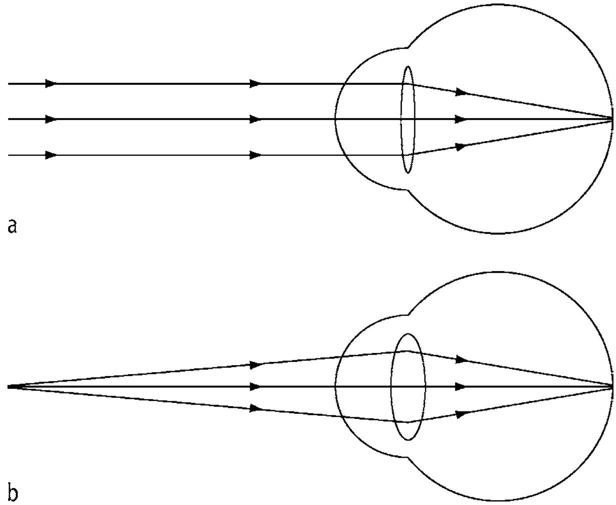


Figure 2.7a,b. Accommodation diagram

loses colour. Other nerve cells depolarise, that is the potential across their cell membranes reduces in response to a stimulus. However, the retinal photoreceptors hyperpolarise, *i.e.* the potential across the cell membrane increases in response to light. This may be due to the fact that the photoreceptor membrane potential in the dark is lower than that of other nerve cells, about 50 mV rather than 70 mV.

In the dark the sodium pores, which are the portions of the receptor membrane with greater permeability to sodium ions, are open and sodium ions flow into the cells and potassium ions flow out. This gives a current called the dark current, which leads to the depolarisation of the receptor and its continued activity. When the visual pigment is bleached in response to light, the sodium pores close, the dark current decreases and the cell is hyperpolarised, leading to a decrease in the release of transmitter, as this is greater at lower rather than higher potentials. The bleaching of the visual pigment leads to the activation of many molecules of an enzyme called transducin, each of which inactivates hundreds of molecules of a chemical called cGMP (cyclic guanosine monophosphate). This leads to the closure of millions of sodium pores, as cGMP is responsible for keeping them open. This cascade response allows a single photon to excite a rod sufficiently to transmit a signal. Rods only respond to changes in illumination in dim light and not in bright light. Their high sensitivity means that the sodium pores are already closed in bright light, making the rods saturated and unable to respond to any further increases in light.

2.4.4 Colour Vision

Visible light consists of electromagnetic radiation with wavelengths between 400 and 700 nm. Both natural light and artificial illumination consist of an approximately even mixture of light energy at different wavelengths, called white light.

When light makes contact with an object, it can be absorbed (and the energy converted to heat), transmitted and/or reflected. In many cases, some light is absorbed and some reflected, with the proportions depending on the wavelengths. A pigment absorbs energy in some wavelengths and reflects others. Pigments appear to be coloured if the refracted energy is in the visible light range. However, the colour perceived depends on the properties of the visual system as well as the wavelengths reflected.

Rods contain rhodopsin, which has a peak sensitivity at about 510 nm in the green part of the spectrum. Rhodopsin is sometimes called visual purple, as it reflects blue and red, making it look purple. There are three different types of cones with peak absorptions at about 430, 530 and 560 nm. These types of cones are therefore referred to as blue, green and red cones. However, these names refer to the peak absorption sensitivities rather than how the pigments appear and monochromatic light with wavelengths of 430, 530 and 560 nm are violet, blue-green and yellow-green rather than blue, green and red. In addition, the three cones have overlapping sensitivity curves. Light at 600 nm will have the greatest response from red cones with their peak sensitivity at 560 nm, but will also have a weaker response from the other two cone types (see Figure 2.8).

The presence of three different types of cones allows colour and brightness to be distinguished, and coloured light to be distinguished from white light. If there were only two types of cones, the monochromatic wavelength that stimulated the two cone systems in the same ratio as white light would be indistinguishable from white light. Therefore, colour-blind people with only two types of cones have a wavelength of light that they cannot distinguish from white light.

The cones function together, whereas the rods work on their own. Although rods and cones may both be functioning in intermediate levels of light intensity, the effects are not combined. People are able to distinguish shapes fairly well in dim light, but are unable to see colours.

Theories of colour vision need to account for the colours of the rainbow, the purples obtained by stimulating the red and blue cones or combining long and short-wavelength light, the pale or desaturated colours obtained from adding white to any other colour and the browns. There are two main theories: the Young-Helmholtz-Maxwell trichromatic theory (see Figure 2.9) and the opponent theory due to Hering (see Figure 2.10). These two theories were originally thought to be contradictory, but it is now realised they are complementary. According to the trichromatic theory, the perception of colour is the result of unequal stimulation of the three types of cones. Light with a broad spectrum of wavelengths, such as light from the sun or a candle, will stimulate all three types of cones approximately equally, giving white. The appearance of white light can also be obtained by using two complementary beams of narrow-band light, such as yellow and blue or red and blue-green which give white when they are mixed together (see Figure 2.9).

The trichromatic theory can predict and explain many, but not all, colour phenomena. For instance, it does not explain why the colour of an object depends on simultaneous and successive colour contrasts, that is, the colours surrounding it and the colours seen immediately before it respectively. According to the opponent colour theory, there are two chromatic visual pairs, comprising yellow and blue,

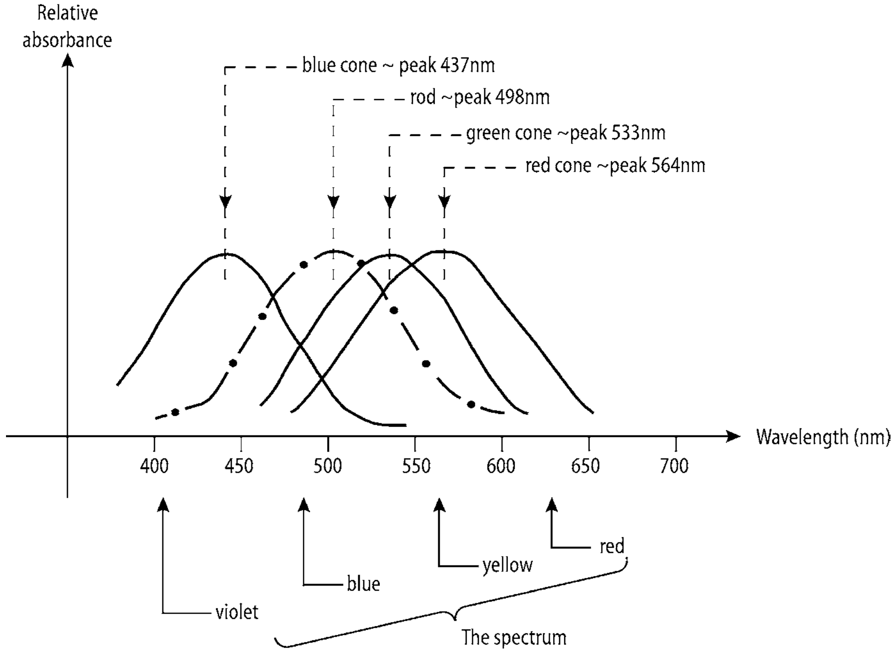


Figure 2.8. Sensitivity of the cones to different wavelengths of light

and red and green and an achromatic pair, white and black. Opponency means that the colours of an opponent chromatic pair cannot be seen together and therefore there are no reddish green or bluish yellow colours (see Figure 2.10).

Some retinal ganglion cells are excited by one colour in the pair and inhibited by the other, for instance, excited by red light and inhibited by green light, and only fire when there is no green light. These cells have a {red-on, green-off} receptive field and do not respond to a combination of red and green lights if the excitatory and inhibitory effects are approximately equal and therefore cancel each other. There are also cells with {green-on, red-off}, {blue-on, yellow-off} and {yellow-on, blue-off} responses. Some cells have centre-surround receptive fields, such as a {red-on centre, green-off surround}.

A {red-on, green-off} cell receives excitatory input from an L (red) cone and inhibitory input from an M (green) cone, whereas a {yellow-on, blue-off} cell receives excitatory input from both M and L cones and inhibitory input from S (blue) cones. When the input is a combination of red and green light, {red, green} cells do not fire due to the opponency of the M and L cone inputs. However, the {yellow-on, blue-off} cells are excited by the M and L cone inputs, whereas the {blue-on, yellow-off} cells are inhibited, so that yellow is transmitted and the combination of green and red is perceived as yellow.

When the input is a combination of yellow and blue light, the yellow light affects the M and L cones and inhibits the {red, green} cells due to opponent input from these cones. The blue light affects the S cones and the {blue, yellow} system is

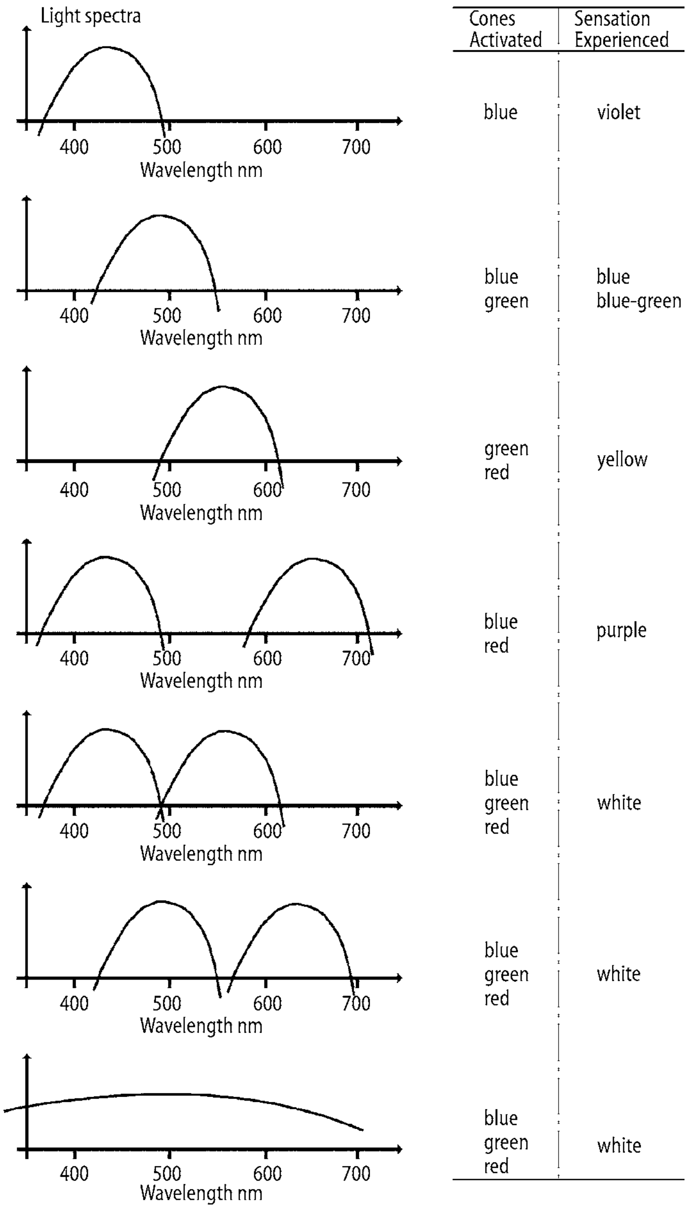


Figure 2.9. Response of the cones to different mixtures of coloured light

inactivated due to the opponent nature of the S and M+L cones. The black channel is also inhibited and therefore white is perceived. When looking at a grey or white background after a bright colour, such as red, people see the contrast colour, in this case green. A bright red flash or prolonged view of red makes the L cones adapt and reduces their sensitivity compared to the M cones. Therefore, subsequent viewing

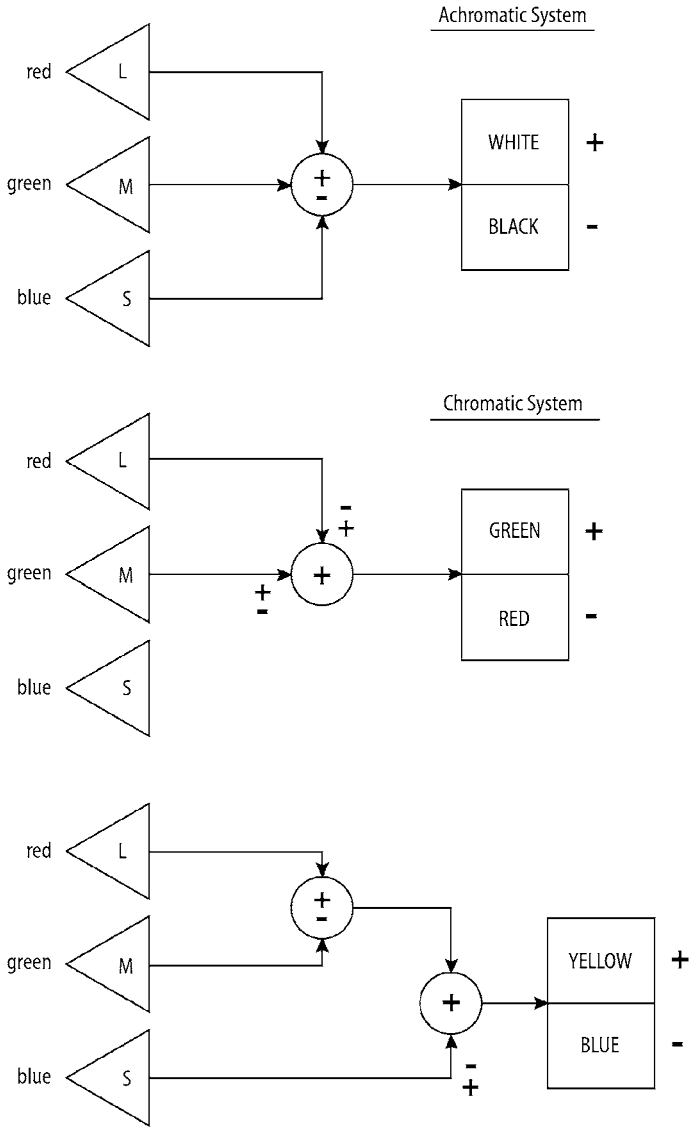


Figure 2.10. Opponent theory of colour vision

of white light with medium and long-wavelengths has a greater effect on the M and L cones, so the green system is activated to a greater extent than the red one. Similar arguments hold for other colour contrasts.

2.4.5 Binocular Vision and Stereopsis

Despite the fact that the image of the three-dimensional world on the retina is two-dimensional, depth vision is important. There are both binocular and monocular cues to depth and distance. Monocular approaches to evaluating depth and distance include parallax or the relative motion of near and far objects when the head is moved from side to side or up and down, perspective and the relative size and position of different objects. However, the most important means of assessing depth, stereopsis, uses the fact that the two eyes receive slightly different images of a scene with depth. Objects whose images are on corresponding points on the two retinas are perceived as being at the same distance, whereas objects which have an outward or inward displacement relative to the retinal image of the point being viewed are perceived as being respectively closer or more distant than this point. About 4–5% of the population without other visual impairments do not have stereopsis, but generally compensate and obtain depth vision by monocular approaches, including parallax, perspective, occlusion and depth perception from movement.

The two retinal images of an object can result in *diplopia* or double vision in the absence of fusion, namely, the perception of the two images as one. It should be noted that fusion, fixation and focus are all different from each other. Fusion is the perception of the two retinal images as one, whereas fixation involves directing the visual axis of one eye at an object to position its image on the fovea and focus the formation of a crisp retinal image. Fusion of the two images into one is only possible when there is approximate correspondence between the retinal locations of the images in the two eyes. The two images do not need to be located at exactly corresponding points, but the image in one eye needs to be located in Panum's fusional area, which is a small area round the corresponding point to the image in the other eye. An image is seen as single when located in the fusional area and double outside it.

A pair of Panum's fusional areas would have similar coordinates on the retina. Corresponding points on the two retinas can be defined either geometrically or functionally in terms of the visual tasks carried out by the two eyes. The term horopter (from the Greek *horos* for boundary or limit and *opter*, observer) is used for the lines or surfaces formed by visual points with images on corresponding retinal points. However it is not uniquely defined.

2.5 Visual Impairment and Assistive Technology

Currently the design of assistive technology generally does not take account of the type of visual impairment of the particular end-user community. Instead, most assistive technology is generally designed for people who are implicitly assumed to be functionally blind, or in some cases, such as screen magnifiers, for people who are generically visually impaired without consideration being given to the specific type of visual impairment. However, only an estimated 3% of blind people are functionally blind and most blind people have some degree of useful vision.

This raises the issue of whether there would be advantages in designing devices for specific groups of blind and visually impaired people. This could lead to the development of devices that are complementary to and take account of residual vision. However, this could have the disadvantage of leading to a very large number of different designs of assistive devices for small numbers of blind and visually impaired people, which may not be technically or financially viable. There may be benefits in developing sets of related devices with different members of the set having features that are appropriate for particular groups of blind and visually impaired people. Thus, this section contains a discussion of the effects of different types of visual impairment, since this may, at some point in the future, have an impact on the design of assistive technology.

2.5.1 Demographics of Visual Impairment

There are an estimated 45 million blind people and 135 million visually impaired people worldwide (WHO 1997a). There are differences both in the percentage of blind people in the population and in the distribution of the different causes of blindness globally. In particular, trachoma, which results from an infection of the conjunctiva by the bacteria *Chlamydia trachomatis* leading to scarring of the cornea, is responsible for 18–24% of blindness globally (Zhang *et al.* 2004), but corneal infection is only responsible for about 2% of cases of registered blindness in the industrialised countries. This is probably because it can be easily avoided by washing the hands and face and treated by antibiotics, though a corneal graft from a donor would be required to restore vision after the cornea has become scarred. Some 90% of blind people live in the 'developing' countries (WHO 1997b).

Ageing frequently has an effect on the visual system, resulting in some degree of visual impairment. For instance, in the UK 82% of blind and visually impaired people are over the age of 50 (TIRESIAS undated) and nearly 73% are over 80 (Bruce *et al.* 1991; Walker *et al.* 1992). In the industrialised countries, the percentage of blind and visually impaired people increases by a factor of three for each decade after the age of 40 (Taylor *et al.* 2001). Many types of blindness are much more common in older people. For instance, age-related macular degeneration is responsible for 39% of cases of registered blindness, as well as being a common cause of blindness in older people, with an incidence of about 8% for people over 65. Cataract is also increasingly common in older people, with an incidence rate of 40% for people over the age of 70 years, with an appreciably greater prevalence in women. It is currently the major cause of blindness globally and responsible for about 45% of cases (WHO 1998).

Glaucoma is the second major cause of blindness globally and is responsible for about 15% of cases of registered blindness. It again increases progressively with age from about 1% at 65 years of age to 7% at 80 years. Over half the population with glaucoma worldwide are unaware that they have this condition (Quigley 1996). Retinal conditions contribute to about 9% of cases of registered blindness, with the incidence of retinopathy due to diabetes increasing from 2% at 50 years of age to 7% at 75 years.

2.5.2 Illustrations of Some Types of Visual Impairment

Many people think that blind people have no vision at all and would be quite surprised to see a person travelling with a long cane reading a book or looking at a map. However, this is quite common and many people who are registered blind do have some useable vision. An indication of what visually impaired and blind people with different types of visions impairments experience is given in the following figures. A typical busy road scene (Figure 2.11) was chosen, as it illustrates one of the important barriers faced by visually impaired and blind people, namely crossing the road or, more generally, finding their way about urban environments safely. Crossing a road can be a difficult and dangerous activity, even for people with no vision or hearing impairments, as cars do not necessarily obey speed limits or other road rules and may seem to zoom up out of nowhere. However, sighted and hearing individuals can use information from both these senses to make decisions as to when it is safe to cross and how fast they need to cross. In the case of visually impaired and blind people considerably less information is available. Therefore unless additional environmental cues are provided by assistive technology, such as ultrasonic or laser canes or audio or tactile 'safe to cross' indicators at crossings, it will be very difficult for visually impaired and blind people to cross roads safely.

Currently, only the following two measures of vision and visual impairment are used:

- Visual acuity or the ability to detect fine details and a quantitative measure of this ability to see an in-focus image at a given distance. Visual acuity is expressed as a fraction or a decimal. For instance, 6:30 means that a person requires to stand at 6 m to see objects that a person without a visual impairment can see

Figure 2.11. Traffic scene

at 30 m. Whilst 6:6 vision implies an ability to understand static images from a distance, it does not necessarily imply perfect vision, as it could be combined with colour blindness, *presbyopia* or an inability to track fast moving objects.

- Field of view; which is the part of the visual world that a person can see at any given time. People have a 180° forward facing field of view out of each eye, but only a 140° field of view from both eyes together.

To give a better appreciation of the quality of a person's vision and their type of visual impairment, there is a need for the development of further measures, including of the following:

- Visual interference effects.
- Motion vision.
- Depth and binocular vision.
- Colour vision.

Loss of visual acuity

Loss of clarity of vision can be specifically localised in extent or vary in quality across the complete field of view. Figure 2.12 shows a case where the central vision is very blurred but with the clarity of vision improving in the peripheral regions of the view. Thus, a person with this form of loss of acuity will be aware of cars at the sides, but will not see them clearly even when they are directly in front of them. They may also have difficulties in estimating the width of the road. Elderly people with age-related macular degeneration due to damage to the pigment epithelium experience this type of view. There are two forms of macular degeneration: the dry and wet forms. In the dry form, material is discarded from the cone outer segments and builds up a layer of deposits behind the pigment epithelium, which impedes the diffusion of nutrients and dispersal of water products. In the wet form, fluid invades the pigment epithelium-photoreceptor space, the cone photoreceptors degenerate and, complete loss of retinal function eventually follows. Although the wet form of the condition may be ameliorated by laser coagulation, vision cannot be restored in either the wet or dry case.

Mild global loss of acuity across the whole field of view is shown in Figure 2.13 and a significant loss of global acuity is shown in Figure 2.14.

In Figure 2.13, the cars and the road are visible to a certain extent, but not very clear. Meanwhile in Figure 2.14, it is very difficult to distinguish exactly how many cars there are or even where the far side kerb is. These two views might be typical of those experienced by people with a mild and an advanced cataract condition, respectively. Vision reduction results from the lens becoming opaque.

Causes of cataract include genetic factors, heat, trauma, uncontrolled diabetes mellitus and ageing. Cataract due to ageing is increasing with increased longevity and may occur in the cortex or nucleus of the lens. The mechanisms involved are not fully understood, but thought to involve a decline in the oxidative status of the lens fibres and an aggregation of crystallins which form water-insoluble aggregates. The sequestration of metals such as iron and manganese leads to a brown/black lens (brunescence). Cataract is treated very successfully by removal of the cataractous lens, leaving the posterior capsule intact (extra-capsular extraction), followed by insertion of a substitute acrylic lens of the appropriate refractive power.

Another cause of reduction in visual acuity is *amblyopia*, where the visual acuity of one eye is reduced due to inadequate visual stimulation during early

Figure 2.12. Traffic scene as viewed with localised central loss of acuity

Figure 2.13. Traffic scene as viewed with mild global loss of acuity

childhood, sometimes in response to other visual impairments. Once established this reduction in visual acuity is in general irreversible. Amblyopia is relatively common with an incidence of 2.0–2.5%. The major causes are occlusion (caused by a congenital cataract), squint (*strabismus*) or unequal refraction between the two eyes. There are two main categories of strabismus: inward (*esotropia*) and outward

Figure 2.14. Traffic scene as viewed with severe global loss of acuity

Figure 2.15. Traffic scene as viewed with nystagmus

(*exotropia*) deviation of one eye, with inward deviation more common. A major cause of esotropia is uncorrected *hyperopia* (long sight), as a result of which a child uses accommodation to increase refraction for distance vision. Unfortunately, this also causes convergence so that one eye serves as the viewing eye while the other becomes amblyopic.

A different type of reduced global acuity is seen in Figure 2.15. In this case, the view is seen as uniformly blurred and it is due to involuntary, rhythmical oscillatory movements of one or both eyes (*nystagmus*). Nystagmus affects 0.01–0.02% of the population. It may be hereditary, due to another neurological condition or temporarily induced by alcohol and other drugs. The resulting visual impairment varies greatly. Other conditions which have nystagmus as a symptom include head trauma, which is the most common cause in young people and stroke, which is the most common cause in older people.

Loss of field of view

Figure 2.16 illustrates the portion of the traffic scene that might be seen by a person with 'tunnel vision'. A person with tunnel vision will not be aware of what is in the peripheral field of view and the person will only be able to see the cars that are directly in front of them, but not those at the sides. They may be able to see across the road and therefore know how wide it is but they may find it difficult to know if their exit on the other kerb is free of obstacles. People with a tunnel vision condition will also experience mobility problems with obstacle avoidance becoming a key issue. Visual field reduction and tunnel vision are typical stages in the various types of *retinitis pigmentosa*. This is an inherited condition, with an incidence

Figure 2.16. Portion of the traffic scene that might be seen by a person with 'tunnel vision'

of 0.03%, in which there is progressive loss of rod photoreceptors resulting in night-blindness. The visual field gradually reduces until only the most central field remains. This is the stage of tunnel vision. Visual acuity also eventually declines, reflecting the loss of cones so that eventually only 5% of the photoreceptors remain.

A condition that leads to a progressive loss of visual field and eventually, if untreated, total blindness is that of glaucoma. It is the result of damage to retinal ganglion cell axons due to an inability to tolerate the prevailing intraocular (inside the eye) pressure. In most cases of glaucoma, the intraocular pressure is raised above the accepted range of 12–21 mm Hg due to impaired drainage of aqueous humour through the canal of Schlemm. However, there are cases of glaucoma without raised intraocular pressures. Glaucoma may have an acute onset, when it is accompanied by considerable pain. However, the usual progression is generally much less apparent so that by the time the person is aware of visual field loss, considerable damage to the retina has already occurred. This could be avoided by regular measurement of visual field and intraocular pressure. Treatment to increase the drainage of aqueous humour may be in the form of eye drops, inhibitors to reduce the formation of aqueous humour or anti-cholinesterases. The eye drops used include timolol or carbonic anhydrase, an enzyme that assists the conversion of carbon dioxide and water into carbonic acid Anti-cholinesterases suppress the action of the cholinesterase enzymes which allow neurons using the neurotransmitter acetylcholine to return to their resting state after activation. In some cases laser treatment or drainage surgery will be required.

A significant reduction in information occurs when a person has only half the field of vision in each eye (*hemianopia*), as shown in Figure 2.17. Consequently, they will be aware of the cars at one side of the road, but not at the other and

Figure 2.17. Portion of the traffic scene seen when half the field of vision is lost

may not be aware how wide the road is. Hemianopia is due to damage to the optic pathways in the brain and can result from acquired brain injuries caused by stroke, tumour or trauma. Hemianopia is classified according to which half of the visual field is absent, with the most common type right homonymous hemianopia, that is both eyes only see the right half of the visual field.

Manifest visual interference effects

Sometimes stationary or dynamic effects can cause a reduction in the quality of a person's vision. These kinds of disturbances can be fixed or transient. However, these visual disturbances obstruct the person's view to the front and/or distract them from the activities they are trying to carry out.

One example in Figure 2.18 shows mild blurriness and blots giving localised loss of vision, typically caused by diabetic retinopathy. This impairment is static. Diabetic retinopathy can result from widespread disturbance of the circulation with consequent hypoxia (reduced oxygen supply) due to uncontrolled diabetes. To try and rectify this new blood vessels are produced in the retina, but may leak and cause haemorrhages. The resulting visual impairment is particularly evident in the blue/yellow colour balance due to a disproportionately great effect on the blue/yellow ganglion cells. Treatment is by laser coagulation to reduce the formation of new blood vessels.

Another form of visual disturbance is the transitory experience of a long slow flash of light often appearing as a slow moving and evolving ring of light as shown in Figure 2.19. This tends to be very distracting, so while this long slow lightshow lasts, the person tends to focus on it rather than the scene. Therefore, although in

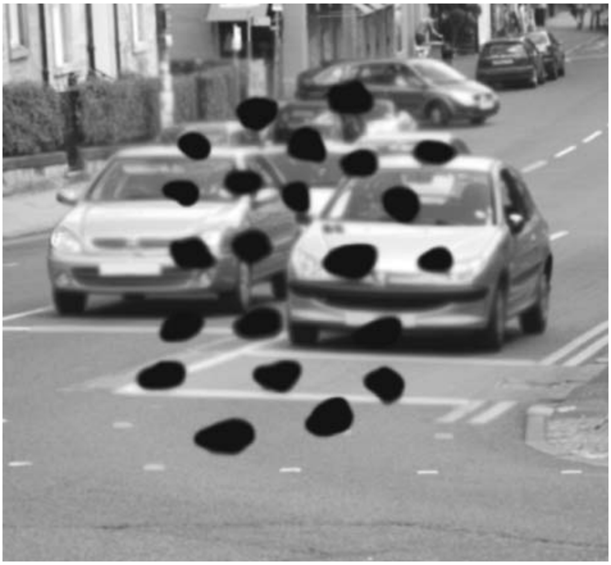


Figure 2.18. Traffic scene as viewed by a person with mild blurriness and blots

Figure 2.19. Traffic scene, as viewed with distracting visual light disturbance

theory the person can see the rest of the scene, in practice they could easily miss seeing some or all of the cars because their attention has been diverted.

These photographs should give some indication of the wide range of visual impairments that people can experience. It should also be noted that vision can vary with the time of day and may be affected by emotional and physical factors such as stress, tiredness, nervousness or unfamiliar surroundings. For instance, many people see much less clearly when stressed or tired. The environment and, in particular, the lighting conditions are important. Most people will see much better in good lighting conditions. However, preferences for different types of lighting often vary with the individual and may also depend on the activities being carried out or other factors. For instance, some people are light-sensitive and may require light sources that provide a particular frequency spectrum.

2.5.3 Further Types of Visual Impairment

Shortsighted people with long eyeballs may develop holes in the retina that could result in detachment of the neural retina from the pigment epithelium. If uncorrected, the photoreceptors that are separated from the choroidal circulation degenerate leading to irreversible degeneration of the post-receptoral neurones. Retinal detachment is responsible for about 2% of cases of registered blindness. Localized detachments can be prevented from worsening by laser treatment in which the detaching retina is spot-welded to the underlying pigment epithelium. Re-attachment may be achieved by surgery to close the holes and reposition the retina.

Loss of the myelin sheath due to inflammation leads to slower conduction along the axons of the visual pathway. In the case of unilateral demyelination, in which neural transmission is slower on one side of the visual pathway than the other, this can lead to double vision and a distortion of vision in which the trajectories of moving objects are seen as curves rather than straight lines. Demyelination can also disrupt saccadic eye movements.

Senile miosis is an involuntary constriction of the pupil with ageing, which reduces retinal illumination leading to a need for brighter ambient lighting. Other factors that reduce retinal illumination due to ageing include increased lens density and a yellowing of the lens, which causes distortion of colour vision. Beyond the age of 60 years, visual acuity itself begins to decline as the visual pigment in the foveal cones decreases, though their number per unit area of retina remains unchanged with age. This increases the visual threshold by a factor of three, thus exacerbating the reduction in light transmission. Therefore, an elderly person viewing the same visual scene as a young person may see it at only one tenth of the brightness. In addition, the ability to detect contrast also diminishes markedly with age due to retinal changes in the cells between the photoreceptors and ganglion cells. However the visual cortex is not affected significantly by ageing.

People with albinism have a lack of pigmentation throughout their body. In particular their hair is absolutely white and their pupils appear as a red disc due to the absence of melanin from the pigment epithelium and choroid. In addition corresponding visual hemifields, such as the left temporal and right nasal fields, are no longer aligned on the same side of the brain with a consequential disruption of binocular vision. Incidentally, Siamese and Burmese cats have the same type of visual impairment.

Blockage of retinal or cerebral vessels, usually by degenerative vascular disease, can result in visual impairment. The retina can be particularly vulnerable to occlusion of the central artery of the optic nerve, which comes directly from the carotid artery and supplies blood to the retina. Occlusion of the posterior cerebral artery can lead to infarction of the visual cortex, though central vision may be preserved. Some localised lesions of the visual cortex can result in *visual agnosia* in which there is a reduced ability to perceive a complete object by combining its individual features, which are generally perceived normally, into complete forms. Other lesions can result in impairments of motion detection (*akinetopsia*).

The term *prosopagnosia* (derived from the Greek *prosopon* for face and *agnosia* for no knowledge) was first used in 1947 for the inability to recognise faces. Most of the documented cases are of acquired prosopagnosia due to head injury, degenerative diseases or infarction (death of cell or living tissue) generally of the posterior zones of the right hemisphere of the brain. A notable case involved a farmer who was no longer able to recognize his family and friends, but could still recognize his cows. However, there are also cases of developmental prosopagnosia, where the onset of the inability to recognise faces occurs prior to the development of full facial recognition abilities in the teenage years. There is still a rather limited understanding of prosopagnosia, including its frequency, particularly with regards to developmental prosopagnosia (FACEBLIND 2001–2006).

2.5.4 Colour Blindness

Colour blindness is an inability to distinguish all the colours that other people can see. The most common types are hereditary photoreceptor conditions due to the presence of a regressive gene on the X chromosome. Therefore, it is generally men who experience colour blindness, whereas women with the regressive gene on one of their X chromosomes are carriers and only actually experience colour blindness if they have the regressive gene on both X chromosomes. However, colour blindness can also be acquired through damage to the retina. The classification of colour blindness is based on the trichromatic theory rather than the opponent theory of colour vision. According to trichromatic theory, colour vision is based on overlap of the absorption spectra of the blue, green and red cone systems in the retina. Different colours are recognised when the different types of cone are stimulated to different extents.

This approach enables colour blindness to be classified either according to the number of functioning cone systems or the difficulty in distinguishing particular colours from each other. The most common types are due to problems with the middle or long-wavelength sensitive cones systems and involve difficulties in discriminating hues or shades of reds, yellows and greens from each other, generally called 'red-green' colour blindness, though the term is an oversimplification. The most common version of this type of colour blindness, which affects about 6% of the male population, is deuteranomaly in which all three cone systems are available, but the medium-wavelength spectrum is shifted towards red, giving a reduced sensitivity to green. Dark green may appear black to people with deuteranomaly in dim light. The intensity of colour vision is not affected. Other types of 'red-green' colour blindness are:

- *Protanomaly*, in which the sensitivity of the long-wavelength pigment is shifted, giving a reduced sensitivity to red and a reduced ability to discriminate colours.
- *Deuteranopia*, in which only two cone systems are used, with the medium wavelength cones absent, giving an inability to distinguish colours in the green-yellow-red section of the spectrum.
- *Protanopia*, in which the long-wavelength sensitive cones are absent, again giving an inability to distinguish colours in the green-yellow-red section of the spectrum.

There are two types of blue-yellow colour blindness:

- *Tritanomaly*, in which the sensitivity of the short-wavelength sensitive cones is affected.
- *Tritanopia*, in which the short-wavelength sensitive cone system is absent, giving an inability to see the bluish-violet end of the spectrum. This is equally common amongst men and women.

Protanomaly, deuteranomaly and tritanomaly are all types of trichromatic (three colour) vision, whereas protanopia, deuteranopia and tritanopia are types of dichromatic (two colour) vision.

Monochromacy or the inability to distinguish any colours is very rare. There are two types:

- Cone monochromacy, in which only one cone system is working, so that there is no colour vision, but there are no other visual impairments.
- Rod monochromacy or maskun, in which the retina does not have any cones, making it difficult to see in light of normal intensity in addition to there being no colour vision.

2.5.5 Corrective Lenses

Corrective lenses are worn on or in front of the eye, generally in the form of spectacles or contact lenses, to treat *myopia*, *hyperopia* (also called *hypermetropia*), *presbyopia* and *astigmatism*. The main aim of corrective glasses is to focus the object image at a single point on the retina. This is called *emmetropia*. Thus, corrective lenses improve vision by improving focus and correcting blur rather than by magnifying images, which would reduce the visual field. However, they can also have small undesirable effects, such as magnification or reduction, distortion, colour fringes and altered depth perception. People with myopia (shortsightedness) use biconcave or diverging lenses of negative power, whereas those with hyperopia (long sightedness) or presbyopia use biconvex or converging lenses of positive power. People with presbyopia in combination with short or longsightedness generally use varifocal or bifocal lenses. People with astigmatism use cylindrical lenses.

Myopia results when the eyeball is too long relative to the focal length of the lens and hyperopia when the eyeball is too short relative to this length. At birth, babies are generally slightly longsighted. The axial length of the human eye increases from 17 mm at birth to 24 mm at maturity with most of the growth occurring during the first three years of life. This causes a reduction in the total refractive power of the eye from +86 D to +55 D, which is necessary for the image to remain sharply focused on the retina.

In people with myopia, the increase in axial length is not fully compensated for by the increase in refraction and the point of focus falls in front of the retina, as shown in Figure 2.20a. The object comes into focus when it is brought closer to the eye without the need to exert accommodative effort. However, biconcave or diverging lenses of negative power are used to correct for this condition, as seen in Figure 2.20b.

The rates of myopia are much higher in countries where people do a lot of close work that strains their eyes, though there is also a hereditary element. Although myopia is normally corrected by wearing distance lenses of appropriate negative power, it has been suggested (MYOPIA undated) that giving children reading glasses with positive power rather distance glasses might stop myopia progressing.

In people with hyperopia there is an insufficient increase in the axial length for the increase in refractive power and therefore the point of focus of the image falls behind the retina, as shown in Figure 2.21a, leading to a blurred image on the retina. People with hyperopia normally see distant objects in focus by

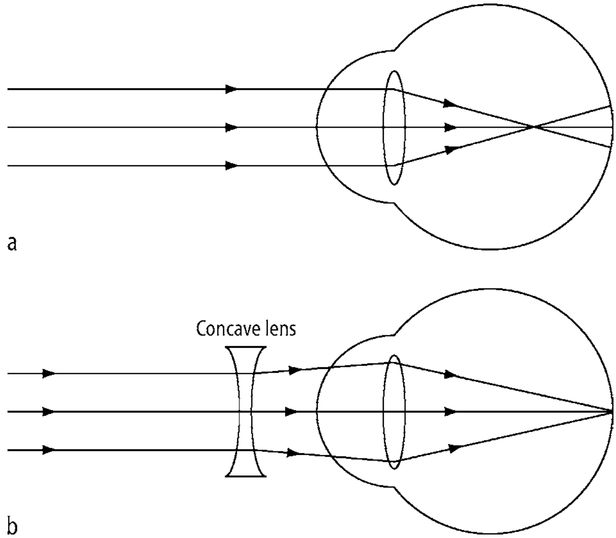


Figure 2.20a,b. Corrective lenses – myopia (shortsightness)

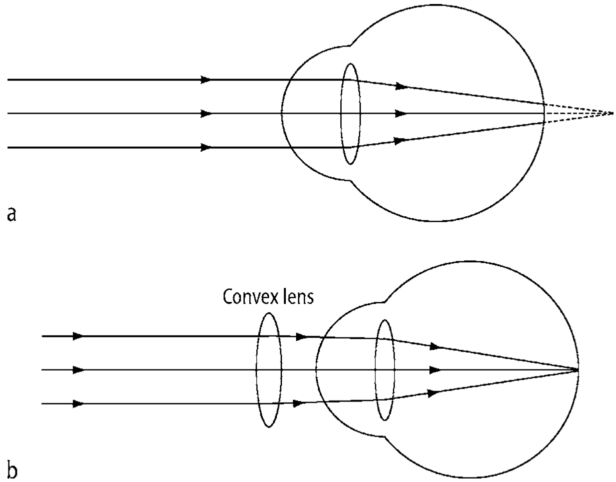


Figure 2.21a,b. Corrective lenses – hyperopia (longsightness)

using increased accommodation, to increase the positive power of the eye. Since there is a limit to the amount of accommodation which decreases with age, there may be insufficient accommodation for viewing near objects. The link between convergence and accommodation (see Section 2.4.2) means that the person may develop strabismus (a squint) in which the eyes are not properly aligned with each other and the direction of gaze. This affects binocular vision and depth perception. People with hyperopia (long sightedness) or presbyopia use biconvex or converging lenses of positive power as shown in Figure 2.21b.

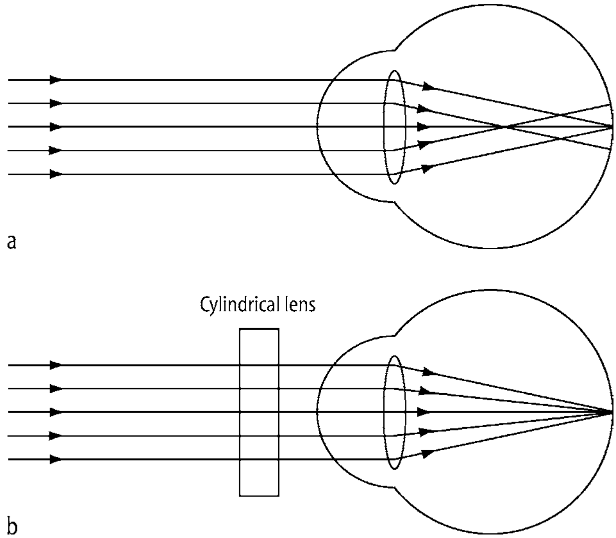


Figure 2.22a,b. Corrective lenses – astigmatism

In astigmatism, the shape of the cornea causes rays of light passing through the lens at different angles to be refracted differently, as shown in Figure 2.22a. As a result there is no level of accommodation that brings all the rays into focus at the same time. This can be rectified by wearing a cylindrical lens of opposite power to the astigmatic error or by wearing a hard contact lens that has a spherical outer surface (see Figure 2.22b). Presbyopia is an inability to focus on nearby objects due to the loss of lens elasticity with age. It affects most people to a certain extent as they get older. The first symptoms of, for instance, difficulty in reading fine print, particularly in poor lighting or eyestrain from reading for an extended period, generally appear at the age of 40–50 years.

Spectacles or glasses are lenses in a frame worn in front of the eye for vision correction or eye protection. The lenses were originally made from glass, but plastic is used to reduce weight and breakages, with scratch-resistant coatings giving a similar scratch resistance to glass. A hydrophobic coating can be used to make the lens easier to clean. Glasses were first used in Northern Italy, probably in the late 1280s. They used convex lenses to correct for presbyopia due to aging. The use of concave lenses to treat myopia is generally ascribed to Nicholas de Cusa. Johannes Kepler's 1602 treatise on optics and astronomy is the first correct published explanation of the use of concave and convex lenses to correct myopia and presbyopia respectively. Benjamin Franklin, who had both myopia and presbyopia, invented bifocals in 1784 to avoid having to use two pairs of glasses. The first lenses for correcting astigmatism were constructed by the astronomer George Airy in 1827.

Contact lenses are corrective (or cosmetic) lenses placed on the cornea of the eye on top of the iris. Contact lenses may be either hard or soft, with soft lenses more commonly used. Soft lenses are often disposable, whereas hard lenses are not. Extended-wear soft lenses can be worn for up to 30 consecutive days before removal, at which point they are discarded. Their extended use is made possible by their high gas permeability. However, they are less often prescribed, as they are more likely to cause problems than other types of contact lenses. Contact lenses are made of polymers containing silicone hydrogel. Many contact lenses are made of water absorbing materials. Toric lenses can be used to correct for astigmatism as well as myopia or hyperopia. They are designed to keep the lens in a stable position despite eye movement. The lens is generally weighted more at the bottom and marked by tiny striations so it can be inserted correctly or designed so that blinking will realign the lens to the correct orientation. Daily disposable lenses do not require cleaning, whereas other types require regular cleaning and disinfection.

Bifocals are generally used by people with presbyopia and either myopia or hyperopia and therefore have corrective lenses of two different lens powers. Since people generally look down when viewing near objects and up when viewing objects at a distance, the reading segment is generally the lower part of the lens and the distance correction for myopia or hyperopia the upper segment. A small reading segment is generally moulded into the lens, with the most popular shape being a D-shaped segment, 28 mm wide. Printed reading matter is generally easily viewed with bifocals, but computer monitors can cause problems, as they are close enough to require corrective lenses, but may be placed so that users have to look up rather than down. The prescription for the reading portion of the lens is generally obtained as a reading addition to the basic prescription.

Trifocal lenses are used by people with presbyopia with two diopters or more of reading addition and have three regions for distance, intermediate (arm's length) and near vision. The reading prescription is obtained as for bifocal lenses and the intermediate prescription is generally obtained from half the reading addition. Therefore a person with a distance prescription of -5 D and a reading addition of $+3$ D, would have a reading portion of $(-5+3)\,\mathrm{D} = -2\,\mathrm{D}$ and an intermediate segment of $(-5 + 3/2)\,\mathrm{D} = -3.5\,\mathrm{D}$. The reading and intermediate segments are generally moulded into the lens. A common style consists of a 28 mm wide D-shaped reading segment with a 7 mm wide intermediate segment inside it. Larger intermediate segments are useful for people who spend considerable time using computers.

Progressive lenses, also called progressive addition, progressive power and varifocal lenses, are gradually replacing trifocal lens. The top of the lens has the distance correction and a gradient of increasing lens powers is added to this to give the reading prescription at the bottom of the lens. By tilting their heads at an appropriate angle, users can obtain the appropriate lens power for each viewing distance. Progressive lenses have the advantage of having a continuous visual field and being more cosmetically attractive than bifocal and trifocal lenses. However, they have regions of distortion and blur away from the optic axis, giving poor visual resolution. As a result, some people are unable to use them.

Corrective lenses have both a spherical and cylindrical component, where the spherical component gives the correction for myopia, hyperopia or presbyopia and acts equally in all directions and the cylindrical component gives the correction for astigmatism (EYEGLASS, undated). People without astigmatism have a zero cylindrical component. The spherical component is given in diopters, with plus values for hyperopia and minus values for myopia. The cylindrical component has both a strength and an axis, as it has a different impact in different directions. The axis specification of the prescription gives the orientation of the axis of the cylindrical correction, which can vary from 1 to 180°. The power of a cylindrical lens varies from zero on its main axis to its maximal value on the axis at right angles to the main axis. The total power of the lens is the sum of the spherical and cylindrical components and therefore equal to the spherical correction on the cylindrical axis and the sum of the spherical and cylindrical corrections on the perpendicular axis.

2.6 Chapter Summary

The chapter has provided an overview of the main structures and functions of the visual system and the different types of visual impairment. In particular, it should be noted that the eye is divided into two chambers by the iris, which regulates the size of the pupil, and consequently the amount of light admitted to the eye. The eyeball has three outer layers, of which the inner sensory layer or retina is particularly important. It contains the photoreceptors: cones, which are responsible for daylight and colour vision, and rods that are used for twilight and night vision. The excitation of the rods and cones and the processing of the resulting signal in the visual cortex of the brain generates an image of the external environment in the brain.

Due to the importance of the photoreceptors in the retina, light from an object entering the eye needs to be focused on the retina. There are two mechanisms involved, eye movements using the three pairs of extraocular muscles to track objects and the optical media of the eye, which refract the incident light rays. The optical media of the eye include the front and back surfaces of the cornea and the biconvex lens, with most of the refraction coming from the front surface of the cornea. Contraction of the ciliary body from which the lens is suspended makes the elastic lens thicker and allows the eye to focus on close up objects. This is called accommodation. The amount of accommodation a person can achieve reduces with age, leading to presbyopia.

Analysis of the spatial form of an object is dependent on the receptive fields of the neurones (nerve cells) in the cortex. Three-dimensional or stereoscopic vision is best obtained from instantaneous comparison of the slightly different images seen by the two eyes, but an impression of depth and perspective can be obtained from monocular vision. Motion detection is performed by complex cells with a directional preference in area V5 of the visual cortex. Colour vision is based on the three different types of cones in the retina, which have different peak spectral absorbencies.

Most assistive technology for blind and visually impaired people is designed for people who are functionally totally blind, though only about 3% of blind people are in this category. The incidence of blindness increases with age and many types of blindness, such as macular degeneration and cataract, are increasingly common in older people. The impact of the different types of visual impairment, including loss of visual acuity and reduced field of view were illustrated by showing how people with these types of visual impairments would perceive a busy road scene.

Questions

Q.1 Briefly define the concept of multisensory perception. What impact does the concept have on the provision of assistive technology for visually impaired and blind people?
Q.2 Draw and label a sketch of a vertical sectional view of the eye.
Q.3 Discuss the role of intraocular eye pressure in eye health.
Q.4 What is the process of accommodation?
Q.5 Give illustrations of three types of visual impairment and discuss what types of safety issues people with these types of visual impairment might experience.
Q.6 Define myopia, hyperopia and astigmatism. Draw diagrams showing the effect of the use of corrective lenses for each condition.

Projects

P.1 Investigate further the concept of multisensory perception. Construct a list of the characteristics of the senses of vision, hearing, touch and smell. Construct a table showing which of these characteristics are unitary and which are complementary. You should indicate any factors which affect the classification. Consider how the sense of smell might be used in assistive technology systems for deafblind people.
P.2 Investigate further the demographics of visual impairment. Construct a table of the inter-relationships between vision, health, age, gender, and the need for assistive technology. Use this table to investigate the types of barriers people in the different categories might experience in carrying out employment, everyday living and recreational activities. Discuss how these barriers might be overcome.
P.3 Use appropriate Web and journal sources to investigate the latest developments in retinal prostheses. Construct a classification of the different research avenues being followed in this work. Give an assessment of the state of the art in each of the categories in your classification.

References

References cited in chapter

Bruce, I.W., McKennell, A.C., and Walker, E.C., 1991, Blind and partially sighted adults in Britain: the RNIB Survey, Vol. 1, HMSO, London
EYEGLASS, undated, http://en.wikipedia.org/wiki.Eyeglass_prescription
FACEBLIND, 2001–2006. http://www.faceblind.org/research/index.html, President and Fellows of Harvard College
Fatt, A., and Weissman, B.A., 1992, Physiology of the Eye: An Introduction to the Vegetative Functions, 2nd ed., Butterworth-Heinemann
Frassinetti, F., Bolognini, N., Bottari, D., Bonora, A., and Ladavas, E., 2005, Audiovisual integration in patients with visual deficit, J. Cognitive Neuroscience, Vol. 17, No. 9, pp 1442–1452
Grossenbacker, P.G., and Lovelace, C.T., 2001, Mechanisms of synesthesia: cognitive and physiological constraints, Trends in Cognitive Science, Vol. 5, No 1, pp 36–41

Lewald, J., and Guski, R., 2003, Cross-modal perceptual integration of apatially and temporally disparate auditory and visual stimuli, Cognitive Brain Research, Vol. 16, pp. 468–478

Morein-Zamir, S, Soto-Faraco, S., and Kingstone, A, 2003, Auditory capture of vision examining temporal ventriloquism, Science Direct, Vol. 17, pp. 154–163

MYOPIA, undated, http://www.preventmyopia.org/

Quigley, H.A., 1996, Number of people with glaucoma worldwide. Br J Ophthalmol., Vol. 80, No. 5, pp. 389–93

Soto-Faraco, S., Lyons, J., Gazzaniga, M., Spence, C., and Kingstone, A., 2002, The ventriloquist in motion: illusory capture of dynamic information across sensory modalities, Cognitive Brain Research, Vol. 14, pp. 139–146

Soto-Faraco, S., Spence, C., and Kingstone, A., 2004a, Congruency effects between auditory and tactile motion: extending the phenomenon of cross-modal dynamic capture, Cognitive, Affective and Behavioral Neuroscience, Vol. 4, No. 2, pp. 208–217

Soto-Faraco, S., Spence, C., Lloyd, D., and Kingston, A., 2004b, Moving multisensory research along: motion perception across sensory modalities, American Psychological Society, Vol. 13, No. 1, pp. 29–32

Spence, S., and Squire, S., 2003, Multisensory integration: maintaining the perception of synchrony, Current Biology, Vol. 13, pp. R519–R521

Spence, S., Baddeley, C., Zampine, M., James, R., and Shore, D.I., 2003, Multisensory temporal order judgments: when two locations are better than one, Perception and Psychophysics, Vol. 65, No. 2, pp. 318–328

Taylor, H.R., and Keeffe, J.E.. 2001 World blindness: a 21st century perspective. Br J Ophthalmol, Vol. 85, No. 3, pp. 261–6

TIRESIAS, undated, http://www.tiresias.org/guidelines/visual.htm

Walker, E.C., Tobin, M.H.J., and McKennell, A.C., 1992, Blind and partially sighted children in Britain: the RNIB Survey, Vol. 2, HMSO, London

Welch, R.B., and Warren, D.H., 1986, Intersensory interactions, In: K.R. Boff, L. Kaufman and J.P. Thomas (eds.), Handbook of Perception and Human Performance, Vol. 1, Sensory Processes and Perception, pp. 25.1–25.36, Wiley, New York

WHO, 1997a, Global initiative for the prevention of avoidable blindness. WHO/PBL/97.61, World Health Organization, Geneva

WHO, 1997b, Global initiative for the prevention of avoidable blindness. WHO/PBL/97.61. World Health Organization, Geneva

WHO, 1998, The World Health report – life in the 21st century: a vision for all, World Health Organization, Geneva, WHO, 1998;47

Zampini, M., Guest, S., Shore, D.I., and Spence, C., 2005, Audio-visual simultaneity judgments, Perception and Psychophysics, Vol. 67, No. 30, pp. 531–544

Zapparoli, G.C., and Reatto, L.L., 1969, The apparent movement between visual and acoustic stimulus and the problem of intermodal relations, Acta Psychol., Vol. 29, pp. 256–267

Zhang, M., Weisser, V.D., Stilla, R., Prather, S.C., and Sathian K., 2004, Multisensory cortical processing of object shape and its relation to mental imagery, Cognitive, Affective and Behavioral Neuroscience, Vol. 4, No. 2, pp. 251–259

Additional reference sources (not cited in chapter)

Berman, E.R., 1991, Biochemistry of the Eye, Plenum Press

Bogdashina, O., 2003, Sensory Perceptual Issues in Autism and Asperger Syndrome, Jessica Kingsley Publishers

Braun, J., Koch, C., and Davis, J.L., 2001, Visual Attention and Cortical Circuits, MIT Press

Buser, P., and Imbert, M. 1992, Vision, (transl. R.H. Kay), MIT

Clement, R.A., 1993, Introduction to Vision Science, Lawrence Erlbaum Associates Ltd

Dowling, J.E., 1987, The Retina, an Approachable Part of the Brain, The Belknap Press of Harvard University Press

Ebenholtz, S.M., 2001, Oculomotor Systems and Perception, Cambridge University Press

Farah, M.J., 2000, The Cognitive Neuroscience of Vision, Blackwell Publishers Ltd
Geldard, F.A., 1972, The Human Senses, 2nd Ed., John Wiley and Sons
Gregory, R.L., and Colman, A.M. (Eds.), 1994, Sensation and Perception, Longman
Hart, W.M. (ed.), 1992, Adlers's Physiology of the Eye Clinical Application, 9th Ed., Mosby Year Book Inc
Hubel, D.H., 1995, Eye, Brain and Vision, Scientific American Library
Kanwisher, N., and Duncan, J., 2004, Functional Neuroimaging of Visual Cognition, Oxford University Press
Levine, M.W., 2000, Fundamentals of Sensation and Perception, 3*rd* edn., Oxford University Press
Matthen, M., 2005, Seeing, Doing and Knowing: A Philosophical Theory of Sense Perception, Clarendon Press.
McIlwain, J.T., 1996, An Introduction to the Biology of Vision, Cambridge University Press
Millar, S., and Al-Attar, Z., 2002, The Müller-Lyer illusion in touch and vision: implications for multisensory processes. Perception and Psychophysics, Vol. 64, No. 3, pp. 353–365
Stein, B.E., and Meredith, M.A., 1993, The Merging of the Sense, 2*nd* Edn., MIT Press
Toyoda, J.-I., Kaneko, M.M.I., and Saito, T., 1999, The Retinal Basis of Vision, Elsevier
Weale, R.A., 1992, The Senescence of Human Vision, Oxford University Press
Zeki, S., 1993, A Vision of the Brain, Blackwell Scientific Publications

3 Sight Measurement

Learning Objectives

The science of the measurement of the properties of sight is an advanced medical and engineering discipline. From a medical viewpoint, sight measurements are a means to diagnosis and route to prosthesis prescription. The engineering challenge in this measurement science is the application of precision engineering to quantify sight parameters using non-invasive methods.

The classical sight parameters are acuity and field of vision. Acuity is quantified over a range of distances (far, intermediate and close) whilst field of vision measures the image spread of the eye. These two test groups commonly support the prescription of lenses to correct for vision deficiencies.

A more modern and recent group of diagnostic tools include those to measure intraocular pressure and eye electrical potential gradients and those tools used to photograph the retina. These all use some advanced engineering methods. Finally, there is now equipment that automates some of the classical sight measurements; some of this equipment is now commonly found in opticians dispensaries.

Clearly there is continuing development of the science and equipment for sight measurement involving a substantial medical engineering input; this is described in this chapter for which the learning objectives are:

- To understand the principles of the measurement of classical and diagnostic sight parameters.
- To appreciate the engineering construction of selected items of sight measurement equipment.
- To review the general trend of current and future research and development for sight parameter measurement equipment.

3.1 Introduction

A wide range of complementary tests is available for assessing an individual's vision. Subjective tests such as visual acuity or visual field measurements rely on a response from the patient to a particular test condition. This response can be a tactile or verbal response. Other tests such as pupil response measurement or ocular pressure measurement are automatically performed with the subject not required to give any verbal or tactile response. An ocular examination will give the ophthalmologist (eye doctor) an indication on the health of the various structures in the eye. Finally, electrodiagnostic techniques are available to obtain objective measurements of function from different stages of the visual pathway.

3.2 Visual Acuity

An assessment of a patient's visual acuity is commonly performed in high street opticians or in ophthalmology departments. The procedure is easy to perform and gives a subjective measure of central vision. The most common method is the Snellen visual acuity test. The test should be performed in a room at least 6 m long although smaller rooms can be used if mirrors are employed to increase the viewing distance. The Snellen chart consists of graded letter sizes. Let us consider the construction of a modern Snellen visual acuity chart. The letter sizes are based on a construction from a 5×4 grid. For example, Figure 3.1 shows the construction of the letter 'U'.

The Snellen distance is defined as the distance, d, at which the basic construction unit of a letter subtends $1'$ of arc. Thus, if the basic construction unit size is denoted $h(m)$ and the distance, $d(m)$, then to subtend $1'$ of arc the relationship needed is

$$\tan(1') = h/d$$

and this re-arranges to give the Snellan distance, d, as

$$d = h/\tan(1')$$

For example, a letter size of 35 mm can be constructed from a basic unit size (h) of 7 mm and this unit would subtend $1'$ of arc at a distance, d, given by

$$d = h/\tan(1') = 0.007\,\text{m} \big/ 2.909 \times 10^{-4} = 24.06\,\text{m}$$

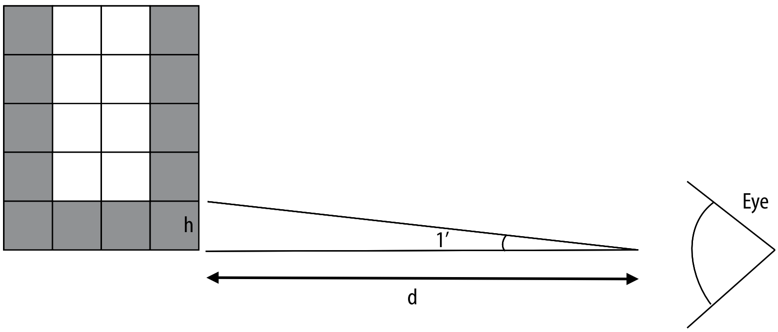


Figure 3.1. Construction of a Snellen letter

V	6/60
T H	6/36
X O A	6/24
H U Y	6/18
V A T U	6/12
O X H W A	6/9
Y U T X I H M	6/6
A X O V T H Y U	6/5

Figure 3.2. Snellen visual acuity chart (not to scale)

Thus, the Snellen size for this letter is therefore 24 m.

The full Snellen chart is shown in Figure 3.2. The letters of the chart are graded in size, based on the Snellen size as described above. For example the top line has a letter of size 87 mm which equates to a Snellen size of 60 m. The second line has letters of Snellen size 36 m and so on down to a Snellen size of 5 m for the bottom line of the chart.

3.2.1 Using the Chart

The patient will normally stand or sit at a viewing distance of 6 m and read the letters. If for example the patient could read down to line 5 without errors then the visual acuity will be recorded as 6/12. If the patient could read to line 7 but got 2 letters wrong on this line then the visual acuity will be recorded as 6/6-2. A recorded acuity of 6/6 is considered normal. In the USA acuity values are quoted in feet rather than metres with 20/20 vision considered as normal visual acuity.

The nurse or optician will often repeat the procedure with the eye covered by a card with a pinhole through which the patient views the chart. This gives a method of assessing the best corrected vision. Viewing the chart through a pinhole eliminates any refractive errors of the peripheral cornea or the crystalline lens of the eye and the acuity measured in this way simulates that with proper glasses in place.

3.2.2 Variations in Measuring Visual Acuity

Another common method of measuring visual acuity and an alternative to the Snellen acuity test is the use of the Bailey–Lovie chart. This chart has the same

Table 3.1. Snellen–LogMAR conversions. The table gives the equivalent MAR and LogMAR values for a particular Snellen acuity

Snellen acuity	MAR	LogMAR
6/5	0.8	−0.08
6/6	1	0
6/9	1.5	0.18
6/12	2	0.30
6/18	3	0.48
6/24	4	0.60
6/36	6	0.78
6/60	10	1.00

number of letters on each line which eliminates any differences due to crowding effects. The letter sizes are also graded in a different way and make use of the minimum angle of resolution (MAR). A Snellen acuity of 6/6 gives a MAR value of 1′ and a Snellen acuity of 6/12 gives a MAR value of 2′ (12/6). These values are converted to a logarithmic scale as shown in Table 3.1.

LogMAR acuity is often preferred in the scientific literature as it is more meaningful to perform statistical analysis on this type of data. In practice a Bailey–Lovie chart will use letter sizes in increments of 0.1 log units and will therefore not correspond exactly to Snellen acuity sizes. For example a LogMAR acuity of 0.5 will correspond to a Snellen value of 6/19 which is not normally available on a Snellen chart.

For patients who are unfamiliar with the alphabet a single optotype Landold C chart can be used This letter C is constructed in such a way that the gap is equivalent to 1 unit of a Snellen acuity letter. The orientation of the letter is varied and can be rotated through 90°, 180° or 270°. The patient is asked to identify where the gap is: top, bottom, right or left (Figure 3.3).

Children will be shown picture acuity cards. This will be a card with a single cartoon picture of a familiar object such as a house, dog or a car. The child is simply asked to identify the picture. The individual cards have the same set of pictures of variable sizes equivalent to the size of the Snellen letters.

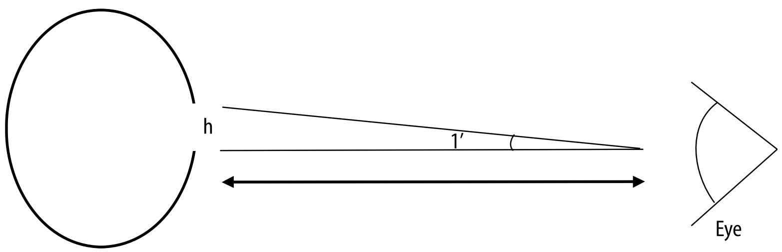


Figure 3.3. Single optotype chart

3.3 Field of Vision Tests

The procedure of testing a patient's field of vision is known as perimetry. The procedure can be performed by an ophthalmologist but is most commonly practiced by a hospital optometrist or a high street optician. A range of methods and instruments are available for perimetry and some of the more common types are described below. In each case the subject is asked to view a central fixation target and different targets presented at various locations in the visual field. The subject responds by pressing a hand-held button if the stimulus is seen.

3.3.1 The Normal Visual Field

If one fixates on a central target then we are able to detect objects in our peripheral vision out to certain limits. A normal subject should be able to detect objects or test stimuli placed within 60° in the superior field, 75° in the inferior field, 100° in the temporal field and 60 degrees in the nasal field. The normal visual field is almost elliptical in shape and this is illustrated in Figure 3.4.

Note that the extent of the normal visual field is dependent on the intensity of the test target. Another important feature of the normal visual field is the location of the physiological blind spot. This is the optic disc area of the retina which is devoid of photoreceptors and is where the nerve fibres bundles pass through to the optic nerve to transmit signals to the visual cortex. The blind spot is located at approximately 15.5° in the temporal field and is a vertical oval approximately 8.5° by 5.5°.

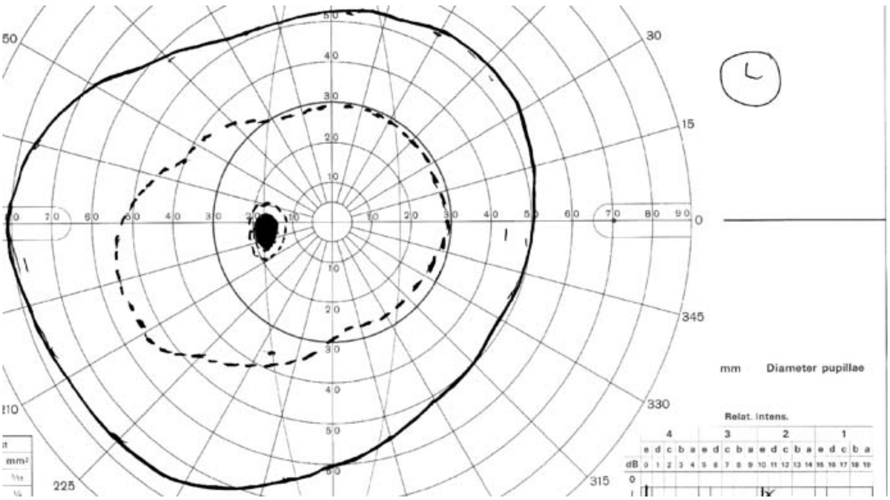


Figure 3.4. Normal visual field plot

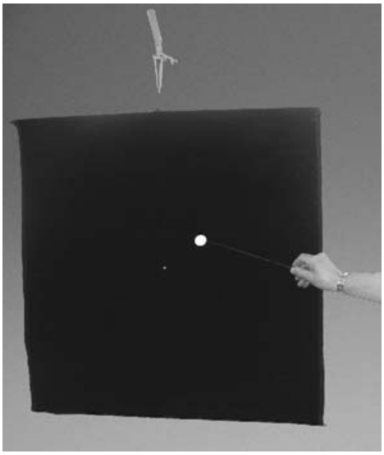


Figure 3.5. Tangent screen

3.3.2 The Tangent Screen

This is a simple and effective method of performing perimetry. A black felt screen (Figure 3.5) of 1 or 2 m^2 is usually used. The centre of the screen has a central fixation mark. Targets are attached to long matt black handles which the operator uses to present the target in the desired location. Targets are usually white in colour and can vary in size from 1–5 mm. The targets are usually black on the reverse side allowing the operator to flip the stimulus quickly to check the reliability of the subject's response.

An isopter defines the boundary between the area where the stimulus is detected and the area outwith the field of vision where the stimulus is not detected. In practice the operator will use large targets to map the blind spot area. The target is moved from non-seeing areas to seeing areas with the patient simply responding that the object has 'gone' or come 'back'. Smaller targets will then be used to assess the visual field. The isopters obtained are expressed as a fraction with the numerator representing the target size and the denominator the screen size. 3/2000 W indicates that a 3 mm white target was used on a 2000 mm screen.

3.3.3 Kinetic Perimetry

Modern perimetry is usually performed with computer controlled automated instruments. The Goldmann perimeter is shown in Figure 3.6. This is a kinetic perimeter which utilises moving stimuli. The principle of operation is similar to that described for the Tangent Screen.

The operator will move stimuli from seeing areas to non-seeing areas with the patient giving a verbal response indicating if the stimulus has 'gone' or come 'back'. The operator will manually document the changes on a chart. The normal visual field from this perimeter was shown in Figure 3.4 and an abnormal example

Figure 3.6. Goldmann perimeter

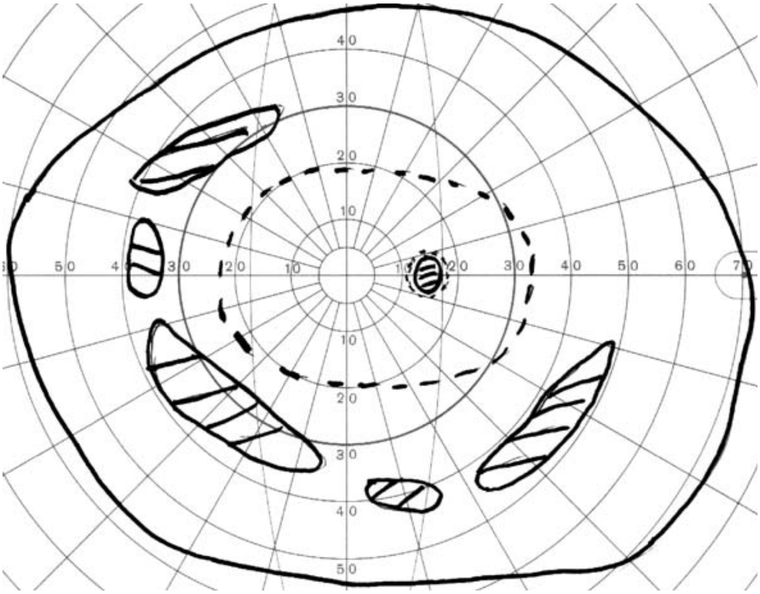


Figure 3.7. Abnormal visual field plot from the Goldmann perimeter

from a patient with the hereditary degenerative condition of *retinitis pigmentosa* is shown in Figure 3.7.

3.3.4 Static Perimetry

The Humphrey perimeter is a common machine used in hospital eye departments and is shown in Figure 3.8. Like most modern perimeters, the device consists of a projection bowl and test stimuli are presented on the surface of the bowl using fibre optics, light emitting diodes or light projections. The bowl design ensures even background illumination of a known intensity which will be consistent between repeated measurements over a period of time. The Humphrey instrument uses light projection to present the stimuli. Protocols vary in complexity but a range of stimulus intensities will be used in each test location in order to determine the threshold for detection. Various algorithms are employed to alter the intensity of subsequent test stimuli based on the subject's previous responses.

Test stimuli will also be placed within the blind spot area to monitor the subject's fixation. Infrared cameras are also used to give the operator an indication of the patient's eye and this is used to monitor fixation throughout the procedure. Throughout the procedure stimuli are presented at pre-determined points in the visual field. If the stimulus is detected then the patient presses a button to indicate the stimulus has been detected.

Figure 3.8. Humphrey perimeter

3.4 Pressure Measurement

Measuring the intraocular pressure (IOP) is important for the diagnosis and monitoring of diseases such as glaucoma. Normal intraocular pressure is in the range 10–21 mm Hg. Direct measurement of IOP is only possible by manometry and this involves penetration of the eye. Tonometry devices give indirect measures of IOP. A number of different types of tonometers are available. The most accurate of the available devices are the applanation tonometers. The Goldmann tonometer is shown in Figure 3.9.

The device is usually used in conjunction with a slit lamp microscope (Figure 3.10) to enable the ophthalmologist to observe the cornea.

A local anaesthetic eye drop is instilled prior to the investigation. The device uses a simple weighted lever system and eccentrically placed weights varied until the applanated area of the cornea is flattened. This is a small corneal area (3 mm) and the test is not uncomfortable for the patient. The weight required to flatten the cornea is directly converted to mm Hg by the device using the equation

$$P = W/A \ ,$$

where P = intraocular pressure, W = weight applied and A = area flattened.

The procedure is repeated several times until two consecutive readings within 0.5 mm Hg are obtained.

The Tonopen (Figure 3.11) is a portable applanation tonometer device which uses strain gauges to measure the pressure with the applanation area equal to 1.5 mm.

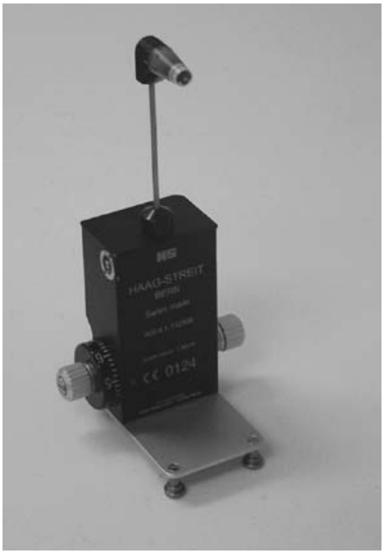


Figure 3.9. Goldmann applanation tonometer

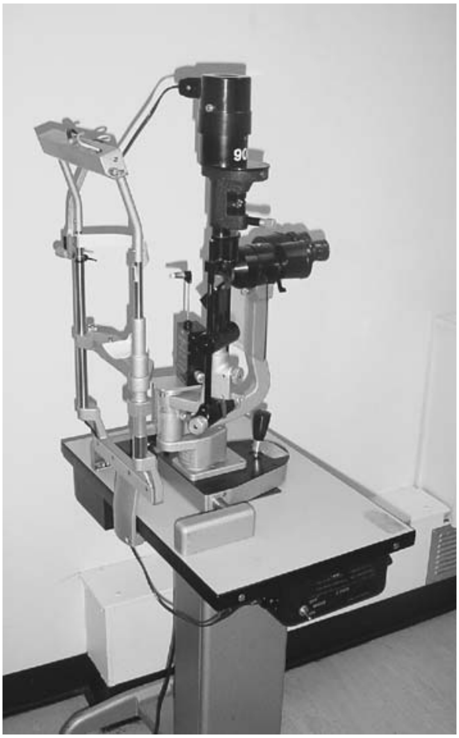


Figure 3.10. Goldmann tonometer in position with the slit lamp

Figure 3.11. Tonopen tonometer

3.5 Biometry

Biometry is the measurement of the length or depth of intra-ocular structures. The test is usually performed before cataract surgery in order to determine the size of artificial lens to implant in the eye. Most devices use a-scan ultrasound.

The cornea is anaesthetised using a topical anaesthetic and the ultrasound probe place on the cornea. The ultrasound frequency is usually 8 MHz and the

Figure 3.12. An a-scan ultrasound biometry device

built-in detector detects the reflections from the various intraocular surfaces. The instrument is shown in Figure 3.12. Modern versions of this instrument employ light rather than sound as the source, making use of laser light and partial coherence interferometry to obtain measurements of the eye's axial length. An instrument that uses a 780 nm laser diode as the source is shown in Figure 3.13 and an example of a typical A-scan is shown in Figure 3.14. If the spike reflections are compared with a cross sectional view of the eye, we can see that we can calculate the axial length of the eye and this information can be used to select which type of lens to implant.

Ultrasound B-scan devices are also available and this enables structures within the eye to be imaged. Higher resolution can be obtained if light rather than sound is used as the source and this is described in Section 3.6.

3.6 Ocular Examination

The ophthalmoscope is a hand-held device which incorporates a light source and a set of magnifying lenses to enable the ophthalmologist to obtain a view of the retina. The device is shown in Figure 3.15. A slit lamp is another examination device that incorporates a moveable light source and a binocular microscope which the ophthalmologist can use to examine the eye.

The device is used by itself to examine the anterior segment of the eye and when it is combined with special lenses it can be used to perform an examination of the posterior segment of the eye. The device was shown previously in Figure 3.11.

The image seen through the ophthalmoscope can also be captured using a digital fundus camera shown in Figure 3.16.

This device produces a hard copy and computer stored version of the retina. The retinal image is shown in Figure 3.17 and has a number of key features such as the

Figure 3.13. Ophthalmic biometry device

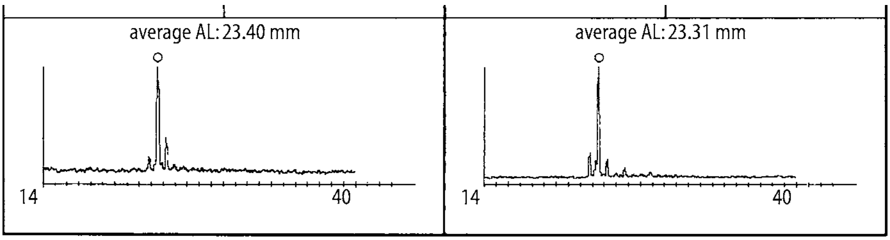


Figure 3.14. Ultrasound a-scan from a normal subject

optic disc, the retinal vasculature and a deeper pigmented area known as the fovea in which the concentration of photoreceptors is greatest and this is the area of fine acuity.

3.7 Optical Coherence Tomography

Optical coherence tomography (OCT) is a new imaging technique which is capable of producing cross sectional images of the retina or cornea with a resolution that surpasses that of conventional imaging techniques.

In principle, OCT is very similar to ultrasonic imaging (Section 3.5), the fundamental difference being that electro-magnetic radiation in the visible or infra-red

Figure 3.15. Ophthalmoscope

Figure 3.16. Digital fundus camera

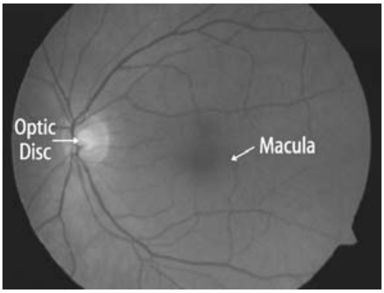


Figure 3.17. Normal retina

portion of the spectrum is used instead of sound waves. An OCT ‘A-scan’ is obtained by shining a beam from a super-luminescent diode onto the tissue to be imaged. As in ultrasound, the principle of echo delay is then employed to calculate the depth of interfaces within the tissue. A series of adjacent A-scans can be combined to form a B-scan, a two-dimensional cross sectional view through the structure. Successive B-scans may be combined to provide tomographic information on the structure.

3.7.1 Echo Delay

In both ultrasound and OCT, when the incident sound or light wave encounters an interface within the tissue, it is partly reflected back towards the emitter and partly transmitted into the tissue. In ultrasound, the depth of an interface is calculated by measuring the time taken for the reflection, or echo, arising from that interface to travel back to the source. The depth of the interface can then be calculated, provided that the speed of sound in the medium is known. Since the speed of light is several magnitudes greater than sound, measuring the echo time delay of a reflected light wave would require ultra-fast resolution, of the order of femtoseconds. In OCT imaging, this problem is overcome by using low coherence interferometry for high resolution measurements.

3.7.2 Low Coherence Interferometry

Figure 3.18 shows a schematic diagram of an optical interferometer. The light source of wavelength λ directs the incident optical wave $E_{\mathrm{i}}(t)$ towards a beamsplitter. At the beamsplitter, one portion of the beam $E_{\mathrm{S}}(t)$ is transmitted into the tissue sample and another portion $E_{\mathrm{R}}(t)$ is reflected towards the reference mirror which is at a known spatial position. The beam which is incident on the tissue sample undergoes partial reflection whenever it encounters a structure or surface within a tissue. Thus, the reflected beam travelling back towards the beamsplitter contains multiple echoes from the interfaces within the tissue. The beam incident on the reference mirror is reflected back towards the beamsplitter. These two reflected beams are recombined in the beamsplitter and the resulting beam $E_{\mathrm{R}}(t) + E_{\mathrm{S}}(t)$ is analysed by the detector.

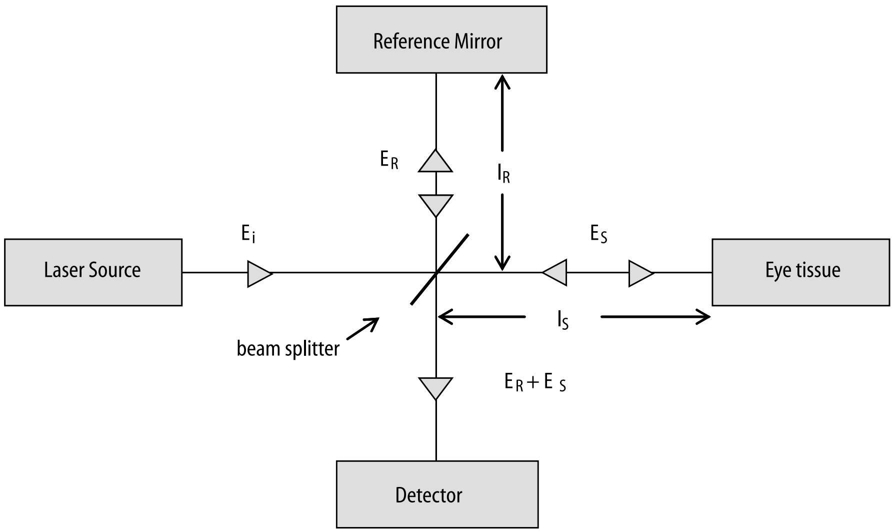


Figure 3.18. Schematic diagram of an optical interferometer

The intensity $I_O(t)$ measured by the detector is proportional to the square of the electromagnetic field. Because of interference effects, the intensity of the output from the interferometer will oscillate as a function of the difference between the path lengths of the reference and specimen beams. From electromagnetic theory, the intensity of the combined beam arriving at the detector is given by

$$I_O(t) = 1/4\,[E_R]^2 + 1/4\,[E_S]^2 + 1/2 E_R E_S \cos\left[2\,[2\pi/\lambda]\,\delta l\right]$$

where l_R is the distance that light travels in the reference path of the interferometer, l_S is the distance that light travels in the measurement path (reflected from the specimen) and $\delta l = l_R - l_S$.

Varying the position of the reference mirror changes the value of l_R and hence δl and will cause the two beams to interfere constructively or destructively. The intensity will oscillate between maximum and minimum each time the path length between reference and measurement arms changes by one optical wavelength (as the position of the reference mirror changes by half a wavelength).

If the light beam is coherent, constructive interference will be observed for a wide range of relative path lengths of the reference and measurement arms. However, in optical imaging it is important to know the position of the structures within the specimen precisely, thus light with a short coherence length must be used. With short coherence length light, constructive interference is seen only when the paths travelled by the reference and measurement beams are nearly equal. If the paths are mismatched by more than the coherence length of the light, no interference effects are observed. The coherence length of light therefore determines the spatial resolution of the imaging system. Since short coherence length light is composed of several frequencies of light, it can be characterized by a frequency or wavelength bandwidth. It can be shown from electromagnetic and optical theory that the range resolution can be related to the bandwidth by the following equation:

$$\Delta L = \left(2 \ln 2/\pi\right) \lambda^2/\Delta\lambda$$

where ΔL is the ranging resolution, λ is the operating wavelength of the light source and $\Delta\lambda$ is the full width half maximum of the spectral bandwidth.

A typical OCT system for ophthalmic applications uses a superluminescent diode operating with a wavelength of around 850 nm and an optical bandwidth of 30 nm. From the above equation this gives an estimated range resolution of approximately 10 μm.

3.7.3 An OCT Scanner

In practice, OCT scanners are built using fibre optics. The light source is a superluminescent diode operating in the near infrared region of the electromagnetic spectrum and a short coherence length of around 10 μm. This is coupled directly into an optical fibre leading to the optical fibre coupler which functions as a beamsplitter. The resultant intereference beam is analysed by a photodiode together with signal-processing electronics and computer data acquisition. The resultant

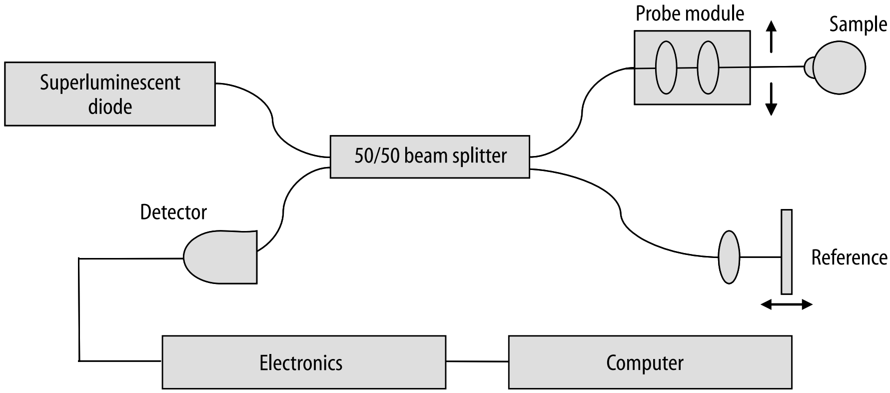


Figure 3.19. Schematic diagram of an OCT scanner

Figure 3.20. OCT scanner

interference data is converted to colour coded maps. A representative schematic diagram of a typical OCT scanner is shown in Figure 3.19.

A commercial OCT scanner is shown in Figure 3.20.

The technique is non-contact and non-invasive and is generally well tolerated by patients. Some sample images taken by an OCT scanner are shown in Figure 3.21.

3.8 Ocular Electrophysiology

Ocular electrophysiology comprises of a range of procedures that enable the visual pathway to be probed in an objective manner. The same equipment can be used for all of the procedures and this is described in the next section.

There are many commercial electrophysiology systems available. The key components of an electrophysiology system are incorporated in the schematic diagram of Figure 3.22. Analog signals are acquired from the subject from electrodes which are attached to the skin using conductive gels. For tests specific to the retina, an additional electrode will be inserted under the eyelid to touch the sclera or cornea. A pre-amplifier will be mounted close to the patient. The purpose of this

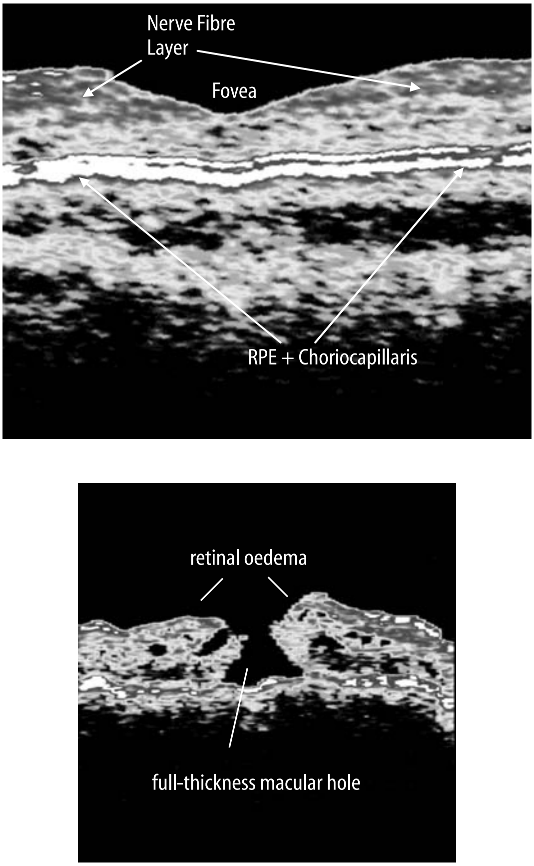


Figure 3.21. Sample OCT images

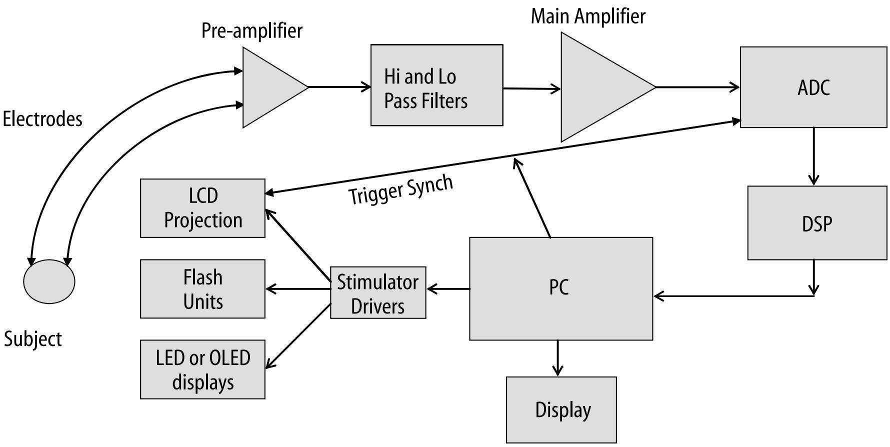


Figure 3.22. Schematic diagram of an evoked potentials system

Figure 3.23. Electrophysiology system in use

differential amplifier is to amplify the signals close to the generator source and this amplifier is usually d.c. coupled, of low gain, typically $\times 10$ and with high common mode rejection ratio (CMMR) to reject common signals to both inputs such as extraneous noise. Following initial amplification, the filter bandwidth will be set to remove d.c. drifts and limit high frequency noise contributions. These filter settings will vary depending on the frequency components of the particular physiological signals to be recovered. The main amplifier is likely to include the filters and this amplifier will have high gain of around $\times 200{,}000$. Once amplified the signals are digitised by an analog to digital convertor (ADC) which will be at least 12 bit resolution. The digital signals are processed by a digital signal processing (DSP) card. Many commercial systems will incorporate the ADC and DSP cards in the personal computer. The DSP card may perform signal conditioning, signal averaging and digital filtering. The personal computer is responsible for controlling the stimulus drivers and for ensuring synchronisation of the analog data with the stimulator. The PC will also display the data and perform the data analysis under user control. A variety of stimulators are available to present flash light or pattern stimulation of the visual system. These stimulators include standard CRT devices, LCD digital projection systems, light emitting diode (LED) displays or organic light emitting diode (OLED) displays).

A commercial electrophysiology system is shown in Figure 3.23.

3.8.1 The Electrooculogram (EOG)

A standing potential exists between the cornea and the retina and this potential is dependent on the level of light or dark adaptation. This potential gives an objective measure of the integrity of the retinal pigment epithelium layer of the retina. It is not possible to obtain a direct measure of this potential but indirect methods can be employed to recover this information.

We can consider this potential to act like a battery or a single dipole in the eye with the cornea being positive with respect to the retina. If we place electrodes either side of the eyes at the outer canthus (temple area) and at the inner canthus then we can measure a potential which is dependent on the direction of gaze.

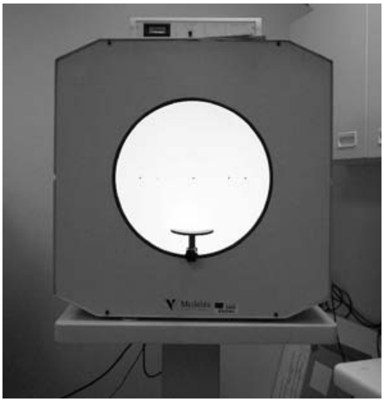


Figure 3.24. Ganzfeld stimulator

Description of test

In order to enable repeatability, reproducibility and comparisons between different clinical and laboratory centres, the EOG procedure has been standardised by the International Society for Clinical Electrophysiology of Vision (ISCEV). The standard sets strict periods of dark and light adaptation and states fixed adaptation light levels.

The standard states that the test should be performed using a Ganzfeld bowl (Figure 3.24).

This ensures an even distribution of illumination across the retina and removes any hotspots that can occur if a standard light source is used. Commercial Ganfeld bowls have red fixation LEDs built in at fixation angles of ±30°. These fixation lights are specifically for the EOG procedure and they can be turned on and off under computer control. At specific intervals (typically every minute) the subject will be asked to perform a series of alternating saccadic eye movements which usually consists of around eight cycles. These saccadic eye movements produce a trace as shown in Figure 3.25.

This trace is an idealised trace which assumes d.c. amplification, perfect eye movements and no impedance drift of the electrodes. This procedure is repeated every 1 min during a period of dark adaptation with the amplitude of the square wave pulse train dependent on the time of dark adaptation. At the end of a 10 min period, the trace amplitude will have fallen to a minimum value. The background illumination is then switched on and this light level is defined as 1.5–3.0 Cd m^{-2} in the international standard and the same procedure is followed for a further 10 min period. By the end of the light adaptation time the square wave pulse train will have shown a significant increase in amplitude.

The Arden ratio

If the amplitude of the square wave is measured and plotted on a graph as a function of time then a graph such as that shown in Figure 3.26 is obtained. This is a graph

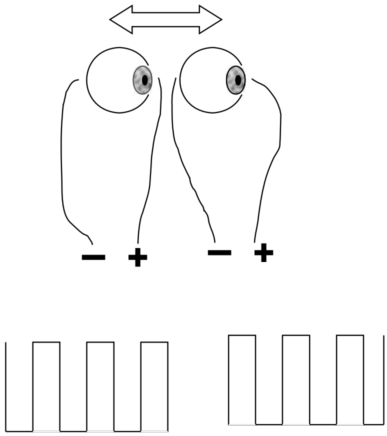


Figure 3.25. Saccadic eye movements produces a square wave output

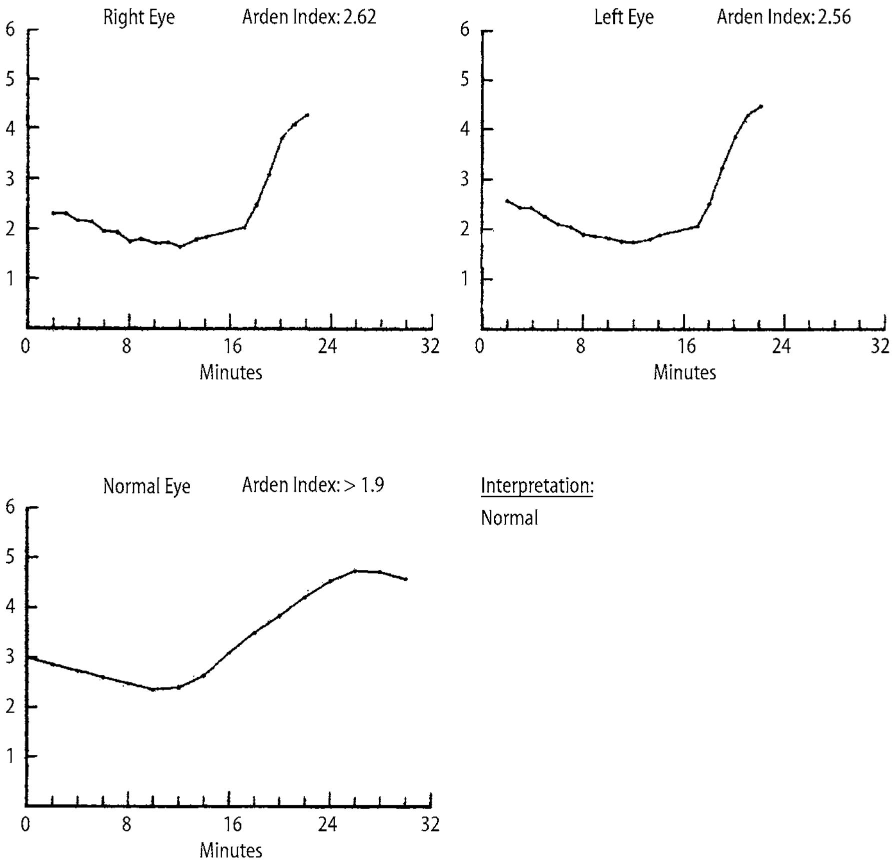


Figure 3.26. EOG Arden graph

for a normal subject and a dark trough is evident after around 10 min followed by a light peak after a further period of light adaptation. If we form the ratio of light peak/dark trough then we obtain a measure of the change in voltage over the period of the test and this ratio is known as the Arden ratio.

3.8.2 The Electroretinogram (ERG)

The electroretinogram was the first physiological signal to be recovered from the human body by two young Scotsmen in 1877. Working independently Dewar and McKendrick recovered an evoked potential to a brief flash of light. This work formed the basis of modern electroretinography which gained popularity in the 1950s with ISCEV formulating a standard for the procedure in 1984.

The stimulus for the full field ERG is the Ganzfeld bowl shown in Figure 3.24. This device gives an even background illumination across the retina with the ability to superimpose short duration full field flashes of light. ISCEV specify the stimulus and recording conditions for the ERG response. A reference skin electrode usually made of Ag/AgCl is placed at the outer canthus. A Ground or indifferent electrode is placed on the forehead or earlobe and the active electrode is placed in contact with the eye. Figure 3.27 shows a range of ERG electrodes that are currently in use. Foils or fibres are placed under the lower eyelid (Figure 3.28) or special contact lenses with electrodes built in can be placed on the cornea.

The ERG response is dependent on a number of variables including the type of electrode used and it is therefore important for each individual laboratory or clinic to establish their own normative data. By appropriate selection of background light levels, stimulus luminance, dark or light adaptation it is possible to obtain a set of responses which give objective and complementary information on the integrity of retinal processing. The examples shown in this section were recorded using the disposable DTL fibre electrode. In the description of the responses a standard flash is defined as a stimulus luminance level of 1.5 to 3.0 $\mathrm{Cd\,m^{-2}\,s}$.

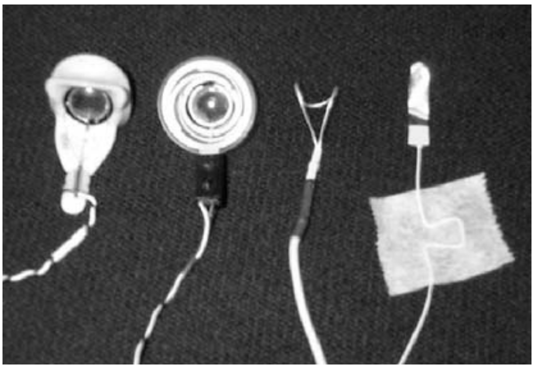


Figure 3.27. ERG electrodes

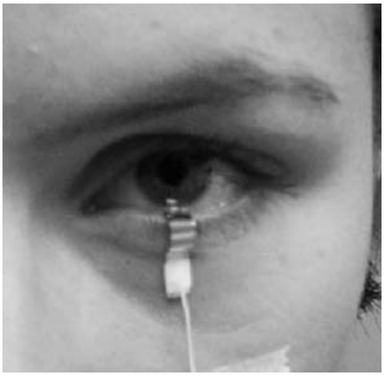


Figure 3.28. Gold foil ERG electrode in place

Response 1: The rod response

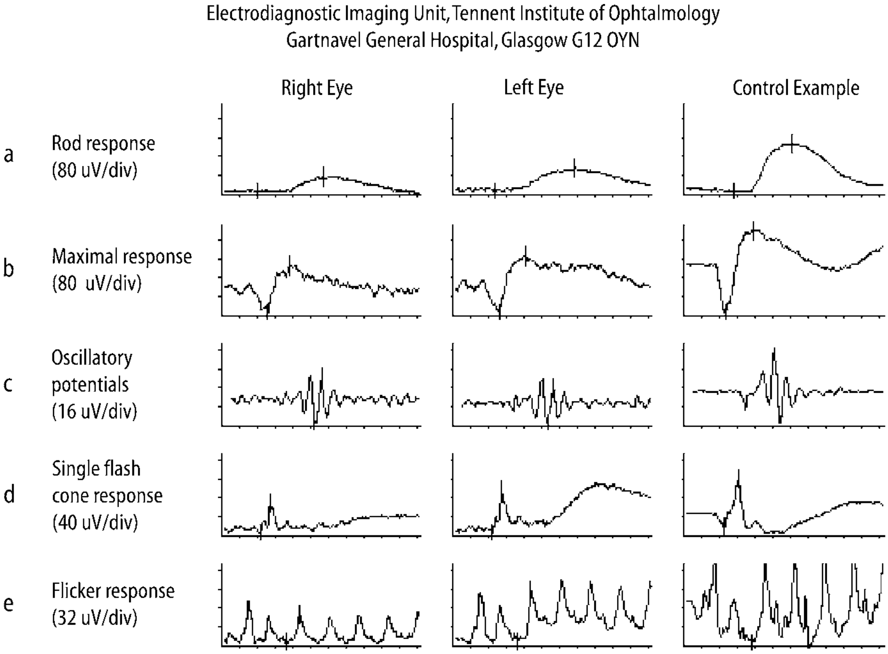
This response is recorded using a dim flash of light in the dark adapted eye. The period of dark adaptation should be at least 20 min. The stimulus intensity should be 2.5 log units below the SF. It is acceptable to perform serial averaging of a number of responses but inter-stimulus duration should not be less than 2 s to avoid light adapting the retina. A typical response is shown in Figure 3.29a. The positive wave is known as the b-wave and is generated by the on-bipolar cells in the retina. As signals are passed to the on-bipolar cells from the rod system, a normal waveform indicates intact rod and on-bipolar cell function. The key measurement is the amplitude from baseline to the peak of the response. Time to peak can also be of interest.

Response 2: The maximal response

This response is also performed on a dark adapted eye. In this case a bright flash of stimulus strength SF is used to evoke a response that is generated by both rod and cone systems. As this is a more intense flash of light inter stimulus duration should not be less than 15 s. An example is shown in Figure 3.29b. The trough is known as the a-wave and is generated by off-bipolar cells with a small contribution directly from the photoreceptors. The positive component is the b-wave and this is mainly generated by off-bipolar cells. The a-wave amplitude is measured from the baseline to the trough and the b-wave amplitude from the a-wave trough to the response peak.

Response 3: Oscillatory potentials

Oscillatory potentials are small oscillations on the rising edge of the b-wave. Stimulation is the same as for the maximal response but in order to emphasize the high frequency oscillations a different amplifier filter bandwidth is used. Instead of 0.5–300 Hz, a restricted bandwidth of 75–300 Hz is recommended. This removes the slow frequency component giving the oscillatory potentials as illustrated in Figure 3.29c.



	Amplitudes			Implicit Times		
	Right Eye	Left Eye	Normal Range	Right Eye	Left Eye	Normal Range
Rod (b-wave)	62	97	72 - 367 uV	79	101	74 - 102 mSec
Maximal (a-ware)	92	103	165 - 291 uV	16	16	15 - 17 mSec
Maximal (b-ware)	L99	238	241 - 709 uV	41	45	34 -59 mSec
oscillatory Pots	37	29	241 - 709 uV	22	22	15 - 20 mSec
Cone (a-wave)	L3	13	17 - 55 uV	8	8	7 - 13 mSec
Cone (b-wave)	63	90	68 - 222 uV	19	19	22 - 31 mSec
Flicker (b-wave)	46	84	25 - 150 uV	19	19	21 - 31 mSec

Interpretation NSR=No significant response

This test was performed in accordance with the 1998 International Standard (IS CEV) for Electroretinography.
Stimulus Intensity=25Cdsm^{-2} Background Luminance=25Cdm^{-2}

Figure 3.29a–e. Standard ERG response

The oscillatory potentials are believed to originate in the inner retina with horizontal and amacrine cells the most likely generators.

Response 4: Cone response

A 10 min period of light adaptation to a background luminance of 17–34 Cd m^{-2} is required before this photopic measurement is performed. In this case the SF is used and an example response is shown in Figure 3.29d. This is a response dominated by the cone pathway with the a-wave generated by the off-bipolar cells and the b-wave the on-bipolar cells. Waveform measurements are the same as in previous examples.

Response 5: Flicker response

A pure response from the cone pathway can be obtained if a fast flickering stimulus is used. In this case, the standard flash is used at a stimulation rate of 30 Hz. The rod system cannot respond at these frequencies therefore the flicker response is a pure cone pathway response. A normal flicker response is shown in Figure 3.29e.

3.8.3 The Pattern Electroretinogram

This response is generated using a reversing checkerboard stimulus. The stimulus is shown in Figure 3.30 and the reversal should take place at a frequency less than 3 Hz. During the reversal, black elements will change to white at the same time as white elements change to black.

Recording electrodes can be the same as in the flash ERG but contact lenses should be avoided as these can interfere with the optics of the eye, degrading the stimulus presentation on the retina. The response is much smaller in amplitude than the standard ERG and is shown in Figure 3.31. The small amplitude of the response necessitates signal averaging with around 200 sweeps required to obtain a reasonable signal-to-noise ratio. The main components are marked on the figure with the N35 and P50 components having similar origins to the flash ERG. A new negative component, the N95 component is believed to originate from the inner retina at the ganglion cell layer. The response is highly dominated by the central retina and there is little contribution from the peripheral retina. This makes the procedure appropriate for objective quantification of macular function.

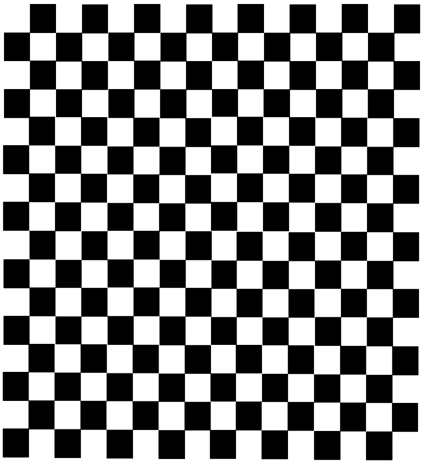


Figure 3.30. Pattern ERG stimulus

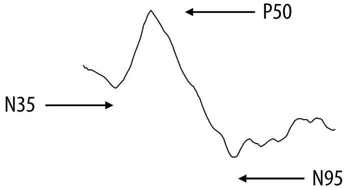


Figure 3.31. Pattern ERG response

3.8.4 The Visual Evoked Cortical Potential

This procedure is used to obtain objective information from the end point of the visual pathway. The stimulus is a reversing checkerboard as previously described and is shown in Figure 3.30. Electrodes are placed at the back of the head over the visual cortex. A reference electrode is placed at a frontal position as indicated in the diagram. Usually, 64 or more serial averages are performed to recover the evoked potential from the background brain activity. The response is shown in Figure 3.32 with the key measurement in this case the time from pattern reversal to the P100 peak.

This is a measure of the transmission time for a signal to be processed by the retina and transmitted down the optic nerve to the visual cortex. The response is highly dominated by central visual function.

3.8.5 Multifocal Electrophysiology

In the past decade or so, a technique of multifocal electrophysiology has revolutionised objective clinical testing of the visual system. Although the standard tests do provide complementary information on the integrity of the various stages of visual processing, they are limited mainly by their global nature. In particular, the ERG is a mass response from the retina which is proportional to the number of photoreceptors in the retina. There is a large variation in the receptor numbers in the normal population and this limits the sensitivity of the test. The test cannot be used to evaluate macular function and is limited to the assessment of generalised retinal dysfunction. Localised retinal dysfunction will therefore not be detected.

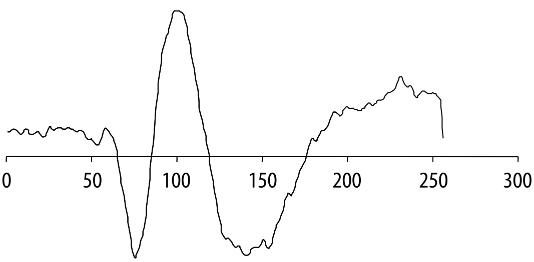


Figure 3.32. Visual evoked cortical potential

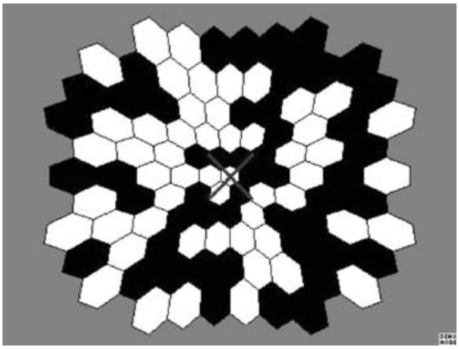


Figure 3.33. Multifocal ERG stimulus

The pattern ERG can give some information on macular function but this test is difficult to perform and gives a single measure of central retinal function. The VECP gives information on central visual function including optic nerve function but gives no information from the peripheral visual field.

The multifocal ERG involves simultaneous stimulation of many retinal sites. A typical multifocal ERG stimulus is shown in Figure 3.33. The stimulus is scaled in this way to take account of the variation in photoreceptor density across the retina.

Focal electroretinography involves the projection of a spot stimulus onto the retina but conventional signal averaging coupled with extremely small signal amplitudes make the test impractical if more than a single area is to be tested. The multifocal technique utilises a special form of pseudo-random binary sequence called m-sequences to stimulate a particular retinal site. These sequences have many useful mathematical properties with the important property being that shifted versions of the same sequence can be run simultaneously and the sequences and therefore the evoked potentials will be orthogonal or truly independent. The sequences are binary and can therefore be in one of two states at any step in time, this is usually black or white. The sequence length is variable but in practice a sequence of length $2^{15}-1$ is the common length for multifocal ERG. The stimulus patches will therefore flicker in a random manner dependent on the control sequence which will run at the stimulation frame rate of 75 Hz. This means that the full sequence for 103 areas for two eyes can be delivered in 32,767 steps or around 7.5 min. This is a considerable saving on the standard ERG recording protocol which can usually take around 1 h due to the light and dark adaptation times. The recording is made from a single active ERG electrode as with the standard ERG and a cross-correlation is performed between the analog data and the particular sequence associated with the area stimulated to obtain a response from that area. The resultant trace array showing 103 small multifocal ERG responses is shown in Figure 3.34.

As with the standard electrophysiology waveforms, there are a number of amplitude and implicit time measurements associated with each waveform. However, the number of parameters for a particular test together with normative data quickly

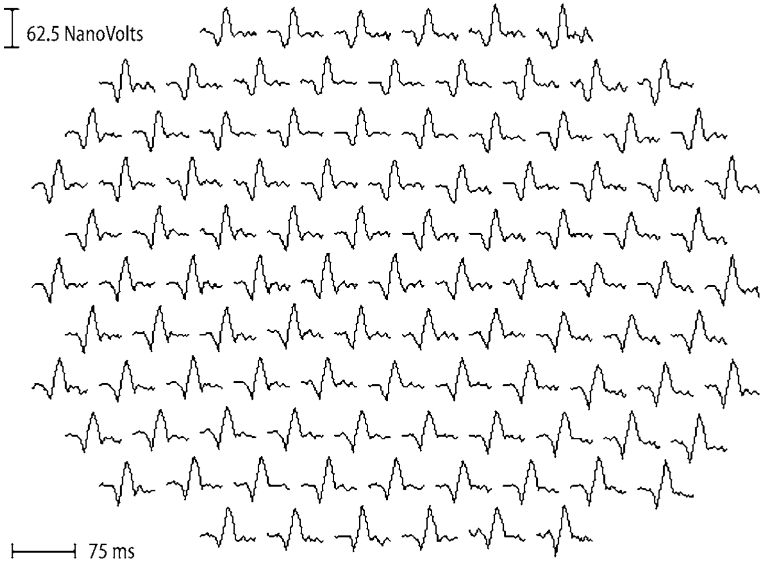


Figure 3.34. Multifocal ERG trace array

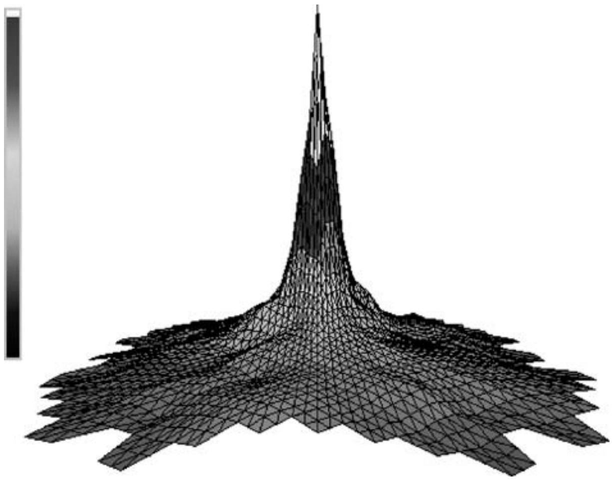


Figure 3.35. Scalar product plot

becomes unmanageable and can be difficult for an ophthalmologist to interpret. Fortunately, another measure known as the scalar product measure is available. This is a measure of deviation from an ideal template waveform and can be used to detect changes in amplitude and implicit time. If the scalar product values are divided by the area stimulated then a response density or scalar product plot can be derived. This plot corresponds to the characteristic hill of vision plots which shows higher function at the central retina where photoreceptor density is greatest. A normal scalar product plot is shown in Figure 3.35. The multifocal ERG

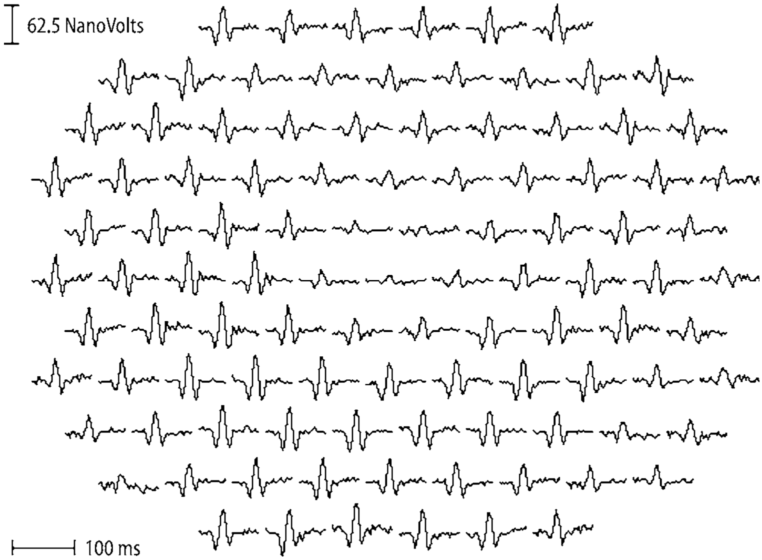


Figure 3.36. Multifocal ERG from a patient with Stargardt's disease (central retinal dysfunction)

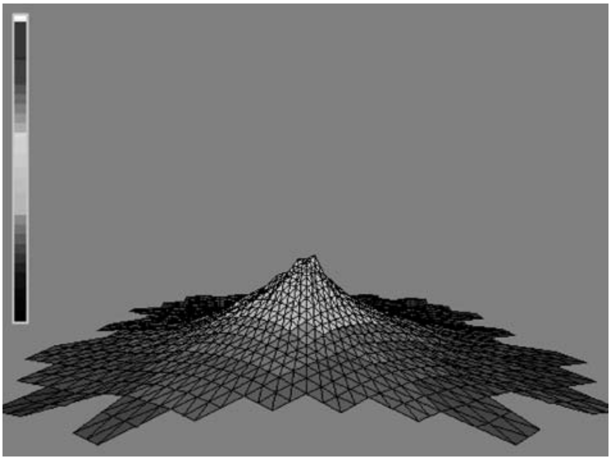


Figure 3.37. Scalar product plot of Stargardt's disease

in a disorder which affects central retinal function such as Stargardt's disease is shown in Figures 3.36 and 3.37.

If a modified version of this stimulus is used for visual evoked cortical potential (VECP) measurement and electrodes placed at the back of the head over the visual cortex it is possible to recover multifocal VECP responses. A typical stimulus scaled for cortical magnification is shown in Figure 3.38 and a multifocal VECP trace array is shown in Figure 3.39.

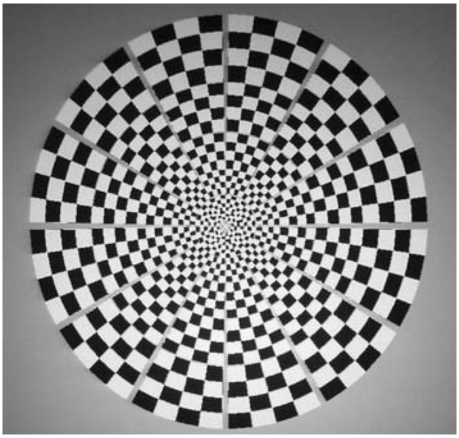


Figure 3.38. Multifocal VECP stimulus

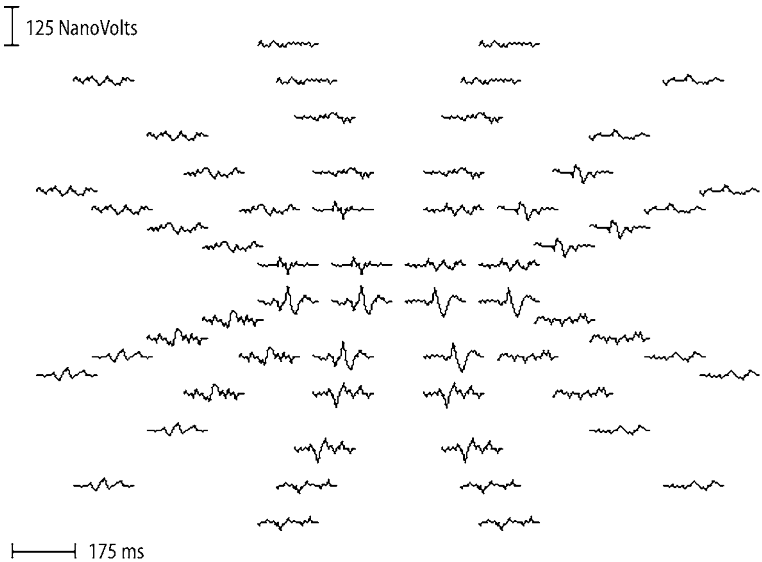


Figure 3.39. Multifocal VECP trace array

The multifocal VECP technique gives clinical electrophysiology the potential of providing objective perimetry. However the technique is less well developed than the multifocal ERG and has several additional difficulties associated with it. VECP responses are variable in shape between individuals and they also vary across the visual field. This is due to different cortical orientations and different dipole source generators. By examining Figure 3.39, it can be seen that responses are smaller in the periphery and there are significant waveform shape changes across the field. Many groups are currently working on addressing these issues and it is hoped that these obstacles will eventually be overcome.

3.9 Chapter Summary

A variety of test procedures and investigations are available to examine the human visual system. A simple visual acuity test (Section 3.2) can quickly tell the investigator if a subject has normal central vision. Normal visual acuity provides valuable information on many key components of the visual system. Good acuity implies that the cornea and lens are clear and free from media opacities which may be caused by cataract or internal haemorrhages. The cornea and lens are also able to focus the image on the retina, the retina is able to process the information and transmit signals down the optic nerves to the visual cortex for interpretation. Good acuity therefore implies an intact functioning central visual system. Complementary to visual acuity, visual field testing will also provide subjective information on the integrity of the peripheral visual system. However, if any of these procedures are abnormal, they are unable to provide information on which part of the visual system an abnormality may be present. In order to localise visual dysfunction, the ophthalmologist will conduct a full clinical examination. This will assess the integrity of the cornea and the lens using the techniques of ophthalmoscopic examination, pressure measurements or biometry as described in Sections 3.5 and 3.6. Many diseases of the retina can be detected by the ophthalmic examination. Abnormalities such as exudates, haemorrhages or abnormal pigmentation may be present which could indicate degenerative disease processes or sight threatening conditions such as diabetic retinopathy. However, the ophthalmic examination does not always reveal the site of dysfunction. This may lie under the retinal surface such as an accumulation of fluid as in macular oedema which can be revealed using new imaging techniques such as OCT (Section 3.7). Structure may also appear to be normal but there may still be cellular dysfunction and defects may be further down the visual pathway such as in optic nerve disease. In these areas objective electrophysiology techniques enable a diagnosis to be reached (Section 3.8).

The combination of subjective testing, ophthalmic examination, electrophysiology and imaging techniques enable powerful diagnosis and monitoring of disease processes or treatment strategies.

Glossary

Acuity: Sharpness and clearness of vision

Electrophysiology: Recording of electrical responses from the muscular or nervous system

Interferometry: Measurement of a length in terms of the wavelength of light

Macula: The central area of the retina – important for central vision

Perimetry: The measurement of visual fields

Retinal pigment epithelium (RPE): An inner retinal cellular layer that absorbs light, provides nutrients and digests ageing visual pigments from the retina

Tonometry: Measurement of intraocular pressure

Questions

Q.1 How can one obtain the best estimate of a subject's visual acuity?
Q.2 How would one perform a visual acuity test on a subject that is unable to read?
Q.3 What would be the disadvantages of performing visual field test binocularly?
Q.4 What level of patient cooperation is required for a visual field test?
Q.5 What technologies are appropriate for examining the structure of the eye?
Q.6 What measurements are made prior to cataract surgery?
Q.7 Draw a schematic diagram of the human retina indicating the key features.
Q.8 Which electrophysiology test gives information on the retinal pigment epithelium layer?
Q.9 Which test can give objective information on the optic nerve function?
Q.10 If a patient had macular dysfunction (reduced central vision), how would the ERG, multifocal ERG, PERG and VECP tests be affected?

Projects

P.1 Construct a simple visual acuity chart and use it to obtain a measure of your own visual acuity. If you wear glasses, try performing the test with and without your glasses. Try repeating the measurement using a pinhole in a card placed in front of your eye.
P.2 Design a simple method to plot your own visual field. Use the technique for other subjects to examine their own visual fields.
P.3 Contact a local ophthalmology department to see the equipment and test procedures described in this chapter.

Reference Sources (not cited in chapter)

Arden, G.B., and Heckenlively, J.R., 1991, *Principles and Practice of Electrophysiology of Vision*, Mosby, St Louis, USA

Celesia, G.C., 2005, Disorders of Visual Processing, Vol. 5, Handbook of Clinical Neurophysiology, Elsevier, Amsterdam, The Netherlands

Edwards, K., and Llewellyn, R., 1988, Optometry, Butterworth and Co., Cambridge

Fishman, G.A., Birch, D.G., Holder, G.E., and Brigell, M.G., 2001, *Electrophysiologic Testing in Disorders of the Retina, Optic Nerve and Visual Pathway*, Ophthalmology Monograph 2, American Academy of Ophthalmology, Singapore

Resources

Standard procedures for all clinical electrophysiology tests are available from the International Society for Clinical Electrophysiology of Vision website, www.iscev.org

4 Haptics as a Substitute for Vision

Learning Objectives

The haptic perceptual system uses a combination of tactile and kinaesthetic information about the environment. Traditionally, this sense has been called touch, but it is nowadays often called haptics to emphasize that not only the sensors in the skin are involved, but also the sensors in muscles, tendons and joints. A related name is active touch. Assistive technology based on haptics has been historically important for the visually impaired and blind person, particularly in the fields of obstacle avoidance for mobility and for accessing print resources. For example, the haptic low technology aid of the long cane made obstacle avoidance when travelling a reality and for printed information, the advent of the haptic medium of Braille made the knowledge and information of printed books accessible.

Despite significant developments with aural interfaces particularly for print, computer systems and mobile telecommunications, haptics will always retain an essential role in the exploration, understanding and use of the physical environment. For this reason it is important to have a basic understanding of the fundamentals and applications of the haptic perceptual system.

The learning objectives for the chapter include:

- Gaining an understanding of the principles and fundamentals of haptics.
- Obtaining an appreciation of the interrelationships between visual and haptic perceptual systems.
- Learning about the different ways in which haptics are used in assistive technology systems.
- Understanding the advantages and also the limitations of haptics in assistive technology systems for the visually impaired or blind person.

4.1 Introduction

The hands are remarkable organs that have had an enormous importance for both the biological and cultural development of human beings. In fact, the hand was in an advanced state long before humans appeared on the scene. Hands have an impressive ability to adapt to different kinds of manipulation tasks, from working with miniature detail to loading heavy objects. It is not as often realised that not only is the hand a marvellous performer, but also it is a highly competent sense organ. Good performance presupposes efficient on-line information about the relationship between the hands and the environment in which they act, a fact that robot developers have had to be quite aware of. There is not yet (ever?) a robot hand that can do all the tasks a human hand can do.

Less well known is the fact that the feet are important as sense organs, too. They can detect information about the ground being walked upon, its slant, material and hardness. This is essential knowledge for safe locomotion. Sighted people are to a large extent informed about these features *via* their eyes, but blind people have mainly to rely on their feet as sensors (Cratty 1971). Other parts of the body surface can also provide information about the environment, for instance, the back and the abdomen, as will be discussed below. The skin is the largest sense organ and very important in many respects.

In textbooks the sensory capacities of the hands are often subsumed under the skin senses. This includes the perception obtained *via* cutaneous sensors, thus not only touch, but also perception of heat, cold and pain. It is certainly true that the tactual information obtained *via* the skin has a basic importance, but this is not the complete story. When the hands are in function, there is a close cooperation between the sensors in the skin and sensors in the muscles, tendons and joints (the kinaesthetic sense), as well as muscles performing the exploration, with all these factors coordinated by the neural system (Cholewiak and Collins 1991; Johnson 2002; Wing *et al.* 1996).

4.1.1 Physiological Basis

Haptic perception is supported by a large variety of receptor and fibre types in the hairless skin and deeper tissues, in all 13 different types of sensory afferents, including four *mechanoreceptive* types providing information about skin deformation and four *proprioceptive* types informative about muscle length, muscle force and joint angle. These two groups of types are the most important for haptic perception, the former type for information *via* the skin, the latter type for information *via* muscles, tendons and joints. Each of the receptor types is specialized for specific kinds of information functions, for example, one type for form and texture perception by being sensitive to edges, corners, points and curvature, another type for perception of hand conformation and of forces parallel to the skin surface. Concerning the proprioceptive information the most important contribution seems to come from the muscle length afferents with some possible contribution from joint afferents. Most of the receptor types are engaged in a majority of the activities of the hand. A thorough analysis of the complex physiological basis and its relation to perception was presented by Johnson (2002).

4.1.2 Passive Touch, Active Touch and Haptics

When touch is described as a skin sense it is often considered as a passive receiver of stimulation from the environment. In opposition to this view, pioneers as Katz (1989) in his seminal work, first published in 1925, and Gibson (1962) emphasized the observers as active explorers of the environment with their hands. The hand is regarded as a perceptual system based on active exploration to collect information, and active touch is considered to be superior to passive touch. To stress the importance of activity the sense is often called "*active touch*" or "*haptics*" (after a Greek word for touching). The latter term will most often be used here, as well as the adjective "*haptic*", sometimes alternating with "*touch*" and "*tactile*" or "*tactual*" in contexts where these terms are traditional. The use of the latter two terms is not consistent, but "tactile" is often used for the physical stimulus (for instance in tactile pictures) and "tactual" for perceptual aspects.

Even if it is has been shown in some contexts that the kinaesthetic component is very useful (see, for instance, the work due to Austin and Sleight 1952, Cronin 1977 and Loo *et al.* 1983), it has sometimes been doubted that activity is a necessary condition for touch to function well. There is experimental evidence that active and passive touch give equivalent results for the perception of texture (Lederman 1981) and small patterns (Vega-Bermudez *et al.* 1991). Magee and Kennedy (1980) found even better results for passive than for active touch in identifying objects in raised line drawings, interpreting the result to depend on favourable conditions for attention to the perceptual task in the passive case. Further experimental analysis is reported by Richardson *et al.* (2004) and a theoretical discussion can be found in Hughes and Jansson (1994). The outcome of this discussion was that active exploration is favourable for the efficiency of haptics for most tasks, but there may be tasks where it is not necessary. Symmons *et al.* (2004) provided an overview of studies concerning the active-passive problem and found the results to a large extent to be task-dependent. Johnson (2002) suggested that passive touch requires more concentration, and that the difference between the two kinds of touch is similar to the difference between situations with dim and bright light in vision.

4.1.3 Exploratory Procedures

The movements performed to collect information *via* the hands are usually not random, but goal-directed. There are specific movements to get specific kinds of information. Lederman and Klatzky (1987) suggested a number of basic exploratory procedures: among others, *lateral motion* for perceiving texture, *pressure* for perceiving hardness, *static contact* for perceiving temperature, *unsupported holding* for perceiving weight, *enclosure* (enclosing the object in a hand or both hands) for perceiving global shape and volume, and *contour following* for perceiving global shape and exact shape. Other exploratory procedures suggested are *wielding* to get information about several properties of an object (Turvey and Carello 1995) and *shaking*, for example, a container with liquid, to be informed about the amount it contains (Jansson *et al.*, 2006). One of the problems with haptic displays, to be discussed below, is that they sometimes require non-natural exploratory procedures that, at least initially, decrease their potential usefulness.

4.2 Vision and Haptics Compared

It can be expected to be an advantage when haptics replaces vision, that the two senses have important properties in common. Here are examples of tasks that both senses can perform, even if haptics usually needs more time for exploration (according to the different procedures mentioned in the above paragraph):

- Find edges separating 3D surfaces.
- Locate objects in relation to observer in near-space (within arms' reach).
- Perceive the size of not too large objects that are possible to explore.
- Perceive the form of not too complicated and large objects.
- Perceive the texture of surfaces (sometimes haptics is better than vision).

Tasks where haptics lags much behind vision or cannot perform the task at all:

- Provide an overview of a scene.
- Perceive 3D space beyond arm's reach.
- Perceive colours.
- Perceive edges in a 2D picture (with no embossment).

Tasks where vision lags behind haptics:

- Perceive the weight of objects (even if vision has capacity to do this in some contexts where the objects are handled by other people; see Runeson and Frykholm 1981 and Jansson 1993).
- Perceive the hardness of surfaces.
- Perceive the temperature of surfaces.

The most important difference is probably the capacity of vision to provide a practically immediate overview of a scene. There are often severe problems in getting an overview of a scene haptically, which may be a laborious and time-consuming task using successive exploratory investigations of the scene. Nevertheless there are also possibilities to identify objects by a "haptic glance", that is a short contact with the object, especially when the observer has hypotheses about what object to expect and the identification is based on local properties such as texture (Klatzky and Lederman 1995).

The difficulty to get en overview from haptics and its lower discriminable capacity has to be considered when constructing graphs. Grids are often used in graphs to facilitate exact reading of values. In tactile graphs grids may be used when the graph is simple, but should be avoided when several lines have to be discriminated because of the risk of grid lines to be perceived as data lines (Lederman and Campbell 1982).

The space covered by the two senses is also a very important difference. Vision allows the sensing of information kilometres away, while haptics is mainly restricted to the space within arm's reach. There are possibilities of extending the haptically reachable space with tools, for instance, the long cane for the visually impaired and, to some extent, *via* information projected onto the skin. Both of these mechanisms are discussed in the sequel. Thus, haptics can provide spatial information, but it has a more limited capacity for overview and 3D space covered than vision.

In spite of these differences it is sometimes possible to use direct translation from visual forms to tactile forms when presenting forms tactually, namely when the figures are not too complicated. However, the more complicated the figure is, the less suitable it may be (Lederman and Campbell 1982).

A very important aspect of choosing tactile symbols is that they should be chosen on the basis of their discriminability tactually. Symbols that are easy to discriminate visually may be quite hard to discriminate when presented for tactual reading (Jansson 1972; Lederman and Kinch 1979).

One possibility to improve the functioning of tactile graphics is to provide redundant information, for instance by making symbols differ in both size and form (Schiff and Isikow 1966).

4.3 The Capacity of Bare Fingers in Real Environments

When investigating the usefulness of technical devices utilizing haptics it is instructive to understand the natural functioning of haptics. It is very often used both for guiding actions and for picking up information, also for sighted people. An everyday example is when you are searching for an object in your pocket or bag without the help of vision. With haptics you can identify the object wanted, grasp it with suitable force and take it out for the use you want. The potential to identify common objects by haptics is quite high, close to perfect within a few seconds (Klatzky *et al.* 1985).

Katz (1989) gave many examples of the capacity of haptics. For instance, it is remarkable that haptics can demonstrate transparency capabilities, as when physicians by palpation of the surface of a body can obtain information about the conditions of an organ under the skin and fat layers. A related property of haptics is remote touching, that is, the experience of a distant object *via* some medium. Physicians can perceive properties of the inner parts of the body also *via* instruments. Visually impaired persons with a long cane can perceive the properties of the ground at the end of the tip of the cane when touching it with the cane. A car driver can feel the goodness of the road *via* the tyres. In the latter example, the observer does not only receive information *via* the hands, but also *via* other parts of the body. Very sensitive parts of the body tactually are the lips and the mouth, which is important for the perception of food. The capacity of the tongue has more recently been used in the context of sensory substitution; see the work due to Bach-y-Rita *et al.* (1998, 2003) and Bach-y-Rita and Kercel (2003). Katz also reported on experiments with arm amputees, asking them to explore surfaces with information from the stump only. Even if the performance was not as good as with the hand, it was remarkably good, which makes understandable that amputees sometimes complain about perception being impoverished with an artificial limb attached.

4.3.1 Visually Impaired People's Use of Haptics Without any Technical Aid

People with severe loss of vision have to rely on other senses, mainly haptics and hearing, in many *everyday tasks*, for instance, find the exact location of a chair to sit on, pick up an object lost on the floor, get the tooth-brush in the bathroom, experience the presence of other persons in the room, and so on. Learning to exploit these alternative possibilities is a lifelong process. Blind people with additional handicaps concerning the other senses have, of course, a much harder time.

Several aspects of visually impaired people's use of haptics are covered in multi-authored books edited by Schiff and Foulke (1982), Heller and Schiff (1991), Heller (2000) and Ballesteros Jiménez and Heller (2004).

In addition to other uses already mentioned here, *social touching*, that is hand-shaking, encouraging pats on the back, hugging, and other more or less intimate contacts between people, may be mentioned. Social touching plays a great role emotionally for most people, handicapped or not, but especially for visually impaired people who are deprived of visual emotional contacts (Thayer 1982). In the case of vulnerable visually impaired and blind people, it is absolutely essential that the person be told who is making the social contact before any physical contact proceeds. This gives the vulnerable person time to decide whether they will accept the social physical contact and in what form they will accept it.

4.3.2 Speech Perceived by Hard-of-hearing People Using Bare Hands

Many people with hearing problems use, especially in communication with other hard-of-hearing people, sign language, that is, they use more or less standardized gestures as a replacement for talking and listening. When communicating with a speaking person they often use visual information from the lip movements of the speaker for understanding (lip reading). One of the problems with lip reading is that it provides ambiguous information about some speech sounds, especially between voiced and voiceless sounds with the same lip movements (b/p, d/t and g/k). A method for a sighted deaf person to enhance the information from visual lip reading is to get tactile information about the vibrations in the vocal tract by holding a hand on the speaker's neck. An elaboration of this method was developed by Öhngren (1992) in cooperation with a totally deaf man. The background of this method was that this man as a boy found out by himself that he could understand speech better if he held a hand on some outer parts of the speaker's vocal tract. As an adult he refined this method further. He even learned to understand a foreign language without ever having heard it spoken.

Lip reading cannot be used by deaf blind people. A remarkable capacity of haptics is revealed by deaf blind people's ability to recognize speech only *via* haptics. Some deaf blind people have been able to perceive the speaker's talking movements by putting fingers on the active parts of the speaker's face (the TADOMA method). The user of this method holds the thumb lightly on the lips and the other fingers on the cheek and neck, thereby obtaining information about movements of the talking person's lips and jaw, airflow and vibrations of the vocal cords, and so being able to perceive speech (Reed *et al.* 1992). This capacity of perceiving speech tactually is obtained only after considerable training, preferably from an early age (Schultz *et al.* 1984). For research on the mechanisms responsible for this tactile capacity an artificial mechanical face capable of producing information similar to that of a talking person was built by Reed *et al.* (1985) and investigations with this device analysed different components involved in understanding speech in this way (Chomsky 1986; Leotta *et al.* 1988; Rabinowitz *et al.* 1990).

Even if it is not the topic of this chapter, it may be mentioned that several tactile aids for the hearing impaired have been developed (Summers 1992).

4.3.3 Natural Capacity of Touch and Evaluation of Technical Aids

As discussed above, haptics has large potentials in many respects when working under natural conditions, including detecting edges, identifying objects and perceiving their texture, hardness, form and weight. It is important to remember these potentials when the efficiency of technical aids is investigated. Likewise, it is important to consider under what conditions the successful performance worked, especially what stimulus properties were available and what exploratory procedures were possible. When a device fails to provide the information wanted important reasons are probably that it is not successfully utilizing the capabilities of haptics, that is, that it does not provide the stimulus properties and exploration possibilities that are essential for the functioning of haptics. As will be discussed below, there are usually many constraints on these basic requirements for the working of present-day haptic displays in comparison with what is naturally available.

4.4 Haptic Low-tech Aids

High-tech is not a necessary requirement for useful technical aids for the visually impaired. This fact is demonstrated by the successful use of the long cane and the guide dog for mobility, Braille for reading and embossed pictures for pictorial information.

4.4.1 The Long Cane

A cane as an aid for the mobility of visually impaired people has been used for centuries. Its basic feature is that it allows the users to extend their arms to reach the ground and objects in front of them. Such an extension of the arm is typical for many tools. The contact between the user and the environment is at the tip of the tool, not at the hand holding the tool. This gives a kind of direct contact with the world simplifying perception. As humans use many tools, the introduction of the cane as a tool is considered to be quite natural.

Even if canes have been used for a long time, they were not used systematically until the twentieth century and especially after the Second World War (Bledsoe 1997). A key feature of the long cane developed at that time was that it was long compared with the canes used previously and that it should mainly be swung over the ground in front of the pedestrian. That makes it possible to preview the ground for about 1 m in front of the user and the space up to the level of the user's waist, which provides information and warnings important for safe walking (Barth and Foulke 1979). In addition to being long, the canes are usually white in order to be more easily observed by other travellers.

In spite of its technical simplicity a long cane can inform its user of many features of the environment, such as location of objects, texture and hardness of the ground, bumps and drop-offs which might cause the traveller to fall. Additionally, it makes possible the detection of street crossings and the direction of kerbs; information that is helpful in the direct guidance of locomotion (Guth and Rieser 1997; Jansson 2000b). The haptic stimulus properties made available *via* the cane are based on vibrations in the cane and force feedback during the handling of the cane. There is also important auditory information, but that is outside the scope of this chapter (Schenkman 1985; Schenkman and Jansson 1986).

There are several practical requirements on such a simple device: for instance, it should conduct vibrations, have a good weight distribution, be lightweight but have sufficient weight to withstand breeze and be strong and durable. For convenience, many people want a collapsible cane, for which there are more requirements. The suitability of the cane tip and the cane grip also needs to be considered, as well as its visibility for other travellers (Farmer and Smith 1997).

It should be noted that it is significant for the usefulness of the long cane that it is handled in a proper way, and that the user has suitable training in its use (Zimmerman and Roman 1997). There are also special programmes developed for preschool children, older persons, as well as for learners with hearing, motor, cognitive and health problems in addition to vision problems; chapters describing these special training programmes can be found in the book edited by Blash *et al.* (1997).

In addition to the long cane, there are many technically simple arrangements that can be useful for the guidance of walking without sight: handrails, traffic sounds and sounding traffic lights (Guth *et al.* 1989; Guth and LaDuke 1994) and aids providing information about veering from an intended path *via* hearing (Gesink *et al.* 1996; Guth *et al.* 1996) or touch (Jansson 1995).

4.4.2 The Guide Dog

At first thought the guide dog for the visually impaired pedestrian may not be considered as a technical aid as the dog is a living creature. However, it may be thought of in that way. The dog and the person form a unit together with the uniting harness. This is a girdle encircling the dog's chest with a handle for the user. With the handle a visually impaired pedestrian both controls the guide dog and receives information about its activities. Haptics is thus important also for guide dog users.

Dogs for guiding visually impaired people were used thousands of years ago, depicted, for instance, on the walls of Pompeii. However, the systematic and widespread use of guide dogs started only after the First World War (Hännestrand 1995; Whitstock *et al.* 1997). As with the long cane, training is an important prerequisite for success. The puppy intended to be a guide dog is carefully selected and trained at special schools, and the guide dog user also has to be trained. Guide dogs are very important for many visually impaired people, but availability and cost restrict their use. It should also be clear that the guide dog, as the long cane, can be helpful for mobility in the near-space, but not for orientation in the larger environment. This information must come from other sources, both from environmental information and from the pedestrian's memory.

4.4.3 Braille

A visually impaired person can read text either by listening to spoken text or using the hands to obtain a tactile equivalent. The former medium is often called talking book; the most common version of the latter medium is Braille. Letters and other symbols in Braille are coded within a six (or sometimes eight) point rectangular matrix that is presented in embossed form and read by the fingers (Foulke 1991). It is an invention that found its form during the early nineteenth century, and it has been a most important reading medium for the visually impaired since then. However, much training is needed to be a fluent Braille reader, which is why it is mainly used by people who are blind from birth or become blind at an early age. The main reason for the difficulties of many older people to learn Braille is that their tactual acuity is not as good as that of a young person (Stevens 1992; Stevens *et al.* 1996). However, the cited problems of older people in learning new skills, slower movements and quicker loss of memory may be unjustified and no obstacle to learning to read Braille in older age (see for example, Millar 1997, pp 239 *ff*).

There is another embossed reading system, Moon, where the letter symbols are similar to ordinary letters and therefore can be read with less training, but this method is not very widespread. However, Moon text has been rendered to be used in connection with simulated tactile diagrams (Hardwick 2004). Technological devices for reading Braille have been developed and these will be discussed below; more discussion can be found in Chapters 12 and 14 of this volume.

A basic prerequisite for reading Braille is that the fingers can discriminate between the point patterns within the matrix. In order to facilitate this, strict rules for the properties of the point matrices have been developed and changes have to be carefully motivated. A thorough analysis of perceptual, linguistic and cognitive aspects of Braille reading was given by Millar (1997).

Many thought that Braille would disappear when the talking book arrived. However, that has not been the case, probably because it has features that the talking book does not have to the same extent, for instance, easier skimming of the text to locate specific parts.

4.4.4 Embossed Pictures

Pictorial information can also be presented in embossed form. Edges and surfaces can be reproduced with embossed lines and patterns of embossed units, respectively. The idea of embossed pictures is quite old; such pictures were available in the eighteenth century (Eriksson 1998). They have not been used as widely as the long cane and Braille, in spite of much effort to make them useful (for example, see Edman 1992). Some kinds of pictures can be read without much difficulty (Kennedy *et al.* 1991) and it has also been reported that blind people can draw pictures (in two dimensions) of three-dimensional objects and scenes (Kennedy 1993). However, more complex embossed pictures are more difficult to perceive haptically, the main problem being to get an overview of the picture. Even if it is possible to get information with a "haptic glance", as discussed above, identification is often obtained only after exploratory movements that can take considerable time and that sometimes do not result in an understanding of the picture. Further, tactile pictures are typically two-dimensional, and to get perception of three-dimensional properties of the scene haptically is usually difficult, even if some possibilities have been reported. Kennedy (1993) described the understanding of perspective by blind people, while Holmes *et al.* (1998) and Jansson and Holmes (2003) reported on the haptic perception of depth in pictures *via* texture gradients, one of the main features of visual 3D perception of pictures.

4.4.4.1 Techniques for Making Embossed Pictures

Originally embossed pictures were "*applications*", that is, the embossments were obtained by gluing things like strings, cloths and other materials onto a paper or board. This method is still used to some extent, but it is an expensive method, as it is necessary to apply all material manually onto each picture. A method making possible the production of copies from *one* such manually made application is *thermoforming*, a method where a model in 3D is copied onto a plastic sheet. In a machine for this procedure the plastic sheet is put onto the model, heated and formed after the model when the air between the sheet and the model is removed. For the production of models, special kits of symbols considered useful for tactile reading have been developed, especially for maps (see, for instance, the work of James and Armstrong 1976 and Barth 1982).

In another method, heating to get embossment is also used – *microcapsule paper* or *swellpaper*. This kind of paper has a layer of microscopic plastic capsules that expand when heated. Black points, lines and patterns written or printed on the paper make the part of paper below them to emboss, while the rest of the paper remains non-embossed at proper heat exposure (see, for instance Edman 1992, pp 82–89). No heating is needed for *raised line drawing boards*. They consist of a thin plastic sheet fastened to some kind of firm backing. The embossment is obtained by writing with, for instance, a ballpoint pen on the sheet. This method is probably the most important method for visually impaired children's drawings (see, for instance, Edman 1992, pp 15–17). Other methods, as well as more details about the methods mentioned here, can be found in works due to Gill (1982) and Edman (1992) pp 13–101.

4.4.4.2 Tactile Maps

Tactile maps are embossed pictures with a special kind of content. They depict geographical conditions, sometimes for teaching in geography, sometimes as a basis for travelling and sometimes both. As embossed pictures in general, they are useful in many contexts, but they also suffer from the difficulty of providing an easy overview. There has been much effort to make them as useful as possible. Overviews can be found in various sources (James and Armstrong 1976; Edman 1992, Chapter 6; Bentzen 1997; Jansson 2003). Different aspects of learning to read a tactile map were investigated by Ungar *et al.* (2004).

4.4.5 The Main Lesson from Low-tech Aids

The main lesson from the success of several low-tech aids is that the usefulness of a technical aid is not a function of its technical complexity. Technically very simple aids, such as the long cane and Braille script, can be very efficient, if they make available the proper information and utilize the natural capacities of the human senses.

4.5 Matrices of Point Stimuli

Even if low-tech aids have been and still are very important for the visually impaired, there have been efforts to replace them or supplement them with more advanced technology. When technological devices for haptics are constructed, efforts have often been made to present an extended surface in contact with the skin, typically providing a matrix of point stimuli. Such a matrix makes it possible to form a pattern within the matrix by dynamically elevating some of the pins above the rest of the matrix. The pins can either be static or vibrating. Many such devices have been built and an overview was given by Kaczmarek and Bach-y-Rita (1995). The same researchers have also developed a device with electrical stimulation to the tongue (Bach-y-Rita *et al.* 2003). Matrices of point stimuli in aids for orientation and mobility, for reading and for tactile pictures (including diagrams and maps) will be discussed below.

4.5.1 Aids for Orientation and Mobility

As discussed above, the long cane is still the most used mobility aid for people with severe visual impairment, and people using guide dogs find usually the dogs very important for their mobility, but it should also be clear that these low-tech aids have constraints. The long cane does not cover the space beyond about 1 m in front of the pedestrian and no space above waist level (when normally used). The guide dog is useful for many aspects of mobility, but it cannot be made responsible for orientation in the larger environment. These constraints have motivated many efforts to construct devices that compensate for the constraints, called ETAs (Electronic Travel Aids). The compensating senses may be hearing or haptics, sometimes in combination. Options based on haptic information will be discussed here; also options based on hearing are discussed by Farmer and Smith (1997) and Jansson (2000b).

An early device intended for the guidance of locomotion was the Electrophthalm (Starkiewicz and Kuliszewski 1963) which reproduced visual information obtained from a camera onto a matrix of pins fastened on the forehead; a later version with a larger matrix (Palacz and Kurcz 1978) is shown in Figure 4.1.

A related device is the Tactile Vision Substitution System (TVSS) with some stationary versions presenting information *via* vibrating pins to, for instance, the back or, concerning mobile versions, *via* electrodes to the chest (Bach-y-Rita 1972); a version of the latter kind is also shown in Figure 4.1.

In these devices the camera on the head picks up environmental visual information that is transformed to a tactile matrix attached to the forehead, the chest or the back. The pattern changes according to motions in the environment and the movements of the pedestrian, as picked up by the video camera. One task is to localise and identify objects at a distance in three-dimensional space, another to guide the user to reach goal objects and to avoid obstacles during walking. These are tasks other than those usually performed by touch, that is, to inform about objects and events in contact with the skin. The skin can be seen to work in a way analogous to the retina of the eye (Bach-y-Rita 1972, p 30). However, for the tactile displays there are apparent limitations in spatial resolution that reduce the possibilities of object identification, as well as collecting distance information. In spite of this it has been reported that users can perceive objects external to the body as localized in the space in front of them (White *et al.* 1970). The reports are anecdotal, however, and there are also conflicting reports of not attaining external localization in spite of long training (Guarniero 1974, 1977).

On the other hand, it has been shown that tactile information provided by the Electrophthalm can guide walking (Jansson 1983). In an experiment the walking person had to make a slalom walk around two poles and towards a goal pole guided only by tactile information from a 12×18 matrix of vibrators on the forehead. The

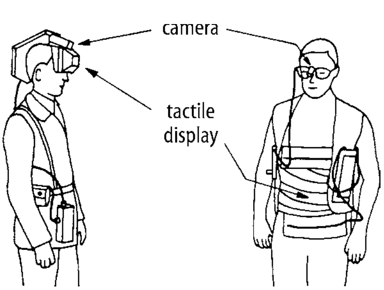


Figure 4.1. Two devices with a matrix of point stimuli fed from a camera, the Electrophthalm (*left*) and a version of the Tactile Visual Substitution System (*right*). (Based on photographs taken by the author)

matrix information is depicted in Figure 4.2 and the resulting movement path is shown in Figure 4.3 together with a visually guided path. Even if the tactually guided walk is not as smooth as the one that is visually guided, it is not too bad.

A related experiment was made with a stationary 20×20 tactile display on the back of the user (Jansson and Brabyn 1981) where the task was to bat a ball running in front of the seated player (Figure 4.4); the tactile information is shown in Figure 4.5. The total time of a game was about 3.5 s, during which time the player had to pick up the tactile information, organize the response and perform appropriate movements. Trained players could perform this task quite well.

Both these experiments demonstrate that tactile information can guide movements in tasks of this kind. However, the visual information transformed to tactual information in these cases is very restricted compared with the information in real

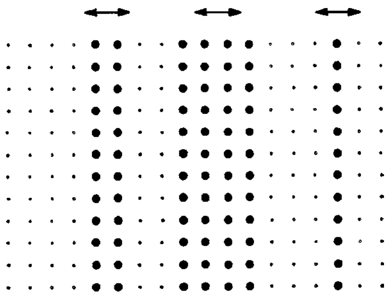


Figure 4.2. Example of a momentary tactile stimulation. From the "slalom walking" study, *three vertical bars* of different widths depicting the slalom poles at different distances and the *arrows* the directions of their motions (Reprinted from Int. J. Neuroscience, 1983, Vol 19, pp. 37–46, Jansson G., Tactile Guidance of Movement, © Gordon and Breach (now part of Taylor and Francis Group) http://www.informaworld.com. Used with permission.)

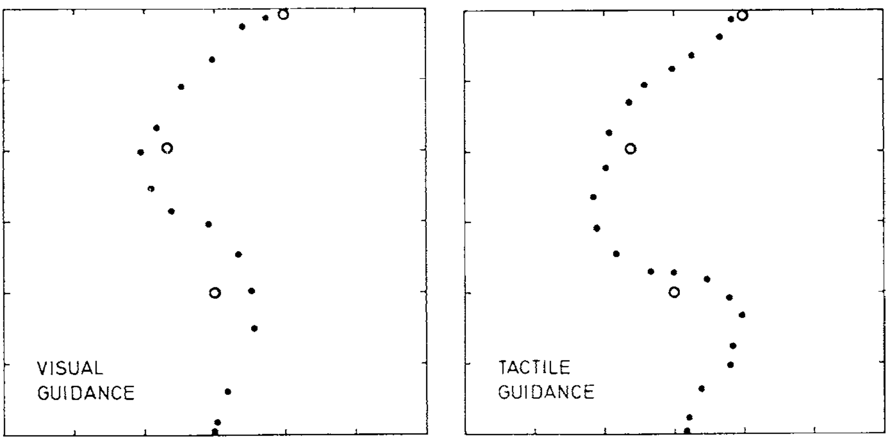


Figure 4.3. Slalom walking movement paths of visual and tactile guidance, respectively (Reprinted from Int. J. Neuroscience, 1983, Vol 19, pp. 37–46, Jansson G., Tactile Guidance of Movement, © Gordon and Breach (now part of Taylor and Francis Group) http://www.informaworld.com. Used with permission.)

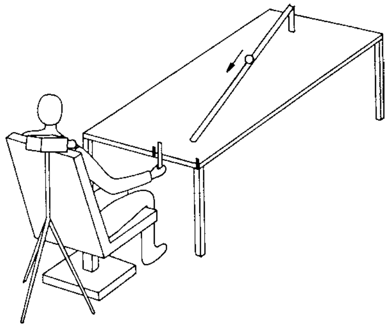


Figure 4.4. Experimental situation in the "batting a ball" study (Jansson and Brabyn 1981)

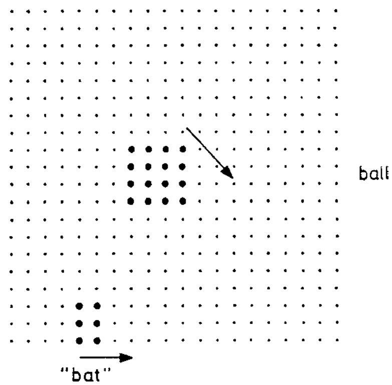


Figure 4.5. Example of a momentary tactile stimulation in the "batting a ball" study. The ball is on an oblique path and "bat" rushing to catch it (Jansson and Brabyn 1981)

environments, black and white representations of a few simple objects. A tactile display of this kind cannot forward sufficiently complex information necessary for guiding movement in ordinary contexts. The Electrophthalm and the TVSS are interesting examples of the capacity of tactile information, but they have not been further developed to practically useable travel aids, probable because of these restricting features.

4.5.2 Aids for Reading Text

Braille is a matrix of point stimuli, usually on paper in embossed form. In the context of computer reading they are presented as protruding pins in the traditional Braille arrangement on a tactile display. There are commercially available devices presenting a row or rows of refreshable Braille cells consisting of pins as an extension of the ordinary computer keyboard (see, for instance, PAP, undated).

A drawback with the use of Braille for texts written with ordinary letters is that it has to be translated into Braille format. A reading aid making direct reading possible is the Optacon (**OP**tical to **TA**ctile **CON**verter). It consists of a camera moved with one hand over a text and transforms the letters into a pattern within a matrix of 5×20 vibrators (its last version, Optacon II) to a pad of a finger on the other hand (Bliss 1978). It can also be used for reading pictures. The Optacon was a success at first and thousands of copies were sold. However, a drawback was that reading the tactile patterns was for most people a slow process and required considerable training, as it was not a natural capacity of touch. With a new technology based on optical reading of a text and transforming it into speech it was much easier to get access to the text, which decreased the demand for the device, and its production was terminated. However, it has also after that been used in research on different aspects of haptic perception.

4.5.3 Aids for Reading Pictures

The Optacon can be used for reading pictures even if this has not been a widespread use of the device. It requires moving the camera over the different parts of the visual figure and perceptually integrating the successive tactual information into a total scene. An alternative to a moving camera is to have a larger display that is explored by moving fingers. As the spatial resolution with present technology is usually a few millimetres, quite a number of pins would be needed in order to simultaneously cover a large scene.

The largest matrix built so far is the Dot Matrix Display DMD 120060, which has a total size of 18×36 cm containing 60×120 pins with a centre distance between the pins of 3 mm (DMD, undated). Such a distance is much larger than the spatial resolution of touch on the finger pad and the display thus does not utilize the full capacity of touch. This display is intended for the visually impaired, but it is quite expensive and only a few devices have been built.

Near the capacity of touch is a display with 20×20 vibrators with a centre distance of 0.4 mm and a total size of 8×8 mm (Pawlik *et al.* 1998). This is a display constructed for basic research on touch and exists so far only in one copy. However, it can give information about the capacity of touch at maximum spatial resolution. A display closer to application with lower spatial resolution (3 mm) but allowing also individual height variation (in 0.1-mm steps to a maximum of 10 mm) of the pins within a 64×64 hexagonal matrix has also been developed (Shinohara *et al.* 1998).

4.6 Computer-based Aids for Graphical Information

There have been several efforts to use computers for the production of tactile maps with Gill (1973) as a pioneer. GISs (geographic information systems) may be a suitable starting point for the content of the map, but they often contain too much information making the tactile maps too cluttered and therefore difficult to read tactually, if all of it is reproduced. Some information is also irrelevant for visually impaired people, and it is often necessary to delete information. Michel (1999) demonstrated that the formats of the systems are differently suitable for this task.

He also suggested that "distortion" of information can be useful in some contexts, for instance by enlarging some areas that are cluttered and decreasing other that contain less information. A system for the production of tactile maps from GIS *via* digital milling to a map put on a touch tablet was developed by Metria (2003).

One effort to increase the efficiency of tactile maps is to combine tactile with co-ordinated auditory information. That can be done by putting a tactile map onto a touch tablet connected to a computer. When the map-reader presses specific points on the tactile map auditory information is obtained, for instance the name of the geographic feature. Pioneering work was made by Parkes (1988) with the NOMAD device. Related work is the interactive auditory learning tool TACTISON (Burger *et al.* 1993) and the dialogue system AUDIO-TOUCH (Lötzsch 1995). A system for audio-tactile graphic information including also an embossment printer has been developed by Viewplus (VIEWPLUS, undated). An increased efficiency in reading tactile maps when they are enhanced with auditory information in this way has been experimentally demonstrated (Holmes *et al.* 1995, 1996). A much larger device of this type, a "Talking kiosk", has been installed at a railway station (Kelly and Schwartz 1999).

4.6.1 Aids for Graphical User Interfaces

Graphical user interfaces (GUIs) are useful for sighted people in interaction with computers, by making it possible to manipulate graphical objects on the screen. Computer users can click on menu items or buttons or move graphical objects on the screen. Usually a mouse is used for these activities, but handling a mouse is difficult without vision. A general effort in the development of aids making it possible for visually impaired computer users to perform these activities in an alternative way was the GUIB project (1995). Its aims were to give the visually impaired access to graphical user interfaces and related textual components, both in the short and long-term. The long-term aim means that directions for the development of an architecture for user interfaces were looked for, directions that could be useful also for visually impaired users by providing non-visual (auditory and/or haptic) options. The efforts included studies of suitable peripherals, filters for extracting information from the graphical environments and screen readers for the presentation of information, as well as development of new pointing devices. Because of the problems to read tactile graphical images a solution was developed where a tactile picture was placed on a touch pad and combined with a written or spoken description in an interactive way. When the computer user touched specific parts of the tactile picture on the pad a verbal description was obtained (cf. the use of the same method for tactile maps, described above).

4.6.2 Tactile Computer Mouse

An ordinary computer mouse usually requires visual feedback, giving the user information about his/her location and movements on a computer screen. A tactile mouse may provide information by a pattern of vibrating or stationary protruding pins within a matrix. One example is the VTPlayer (VTPL, undated) that is a mouse with two 4×4 arrays of pins, each intended for one finger. In addition to the kinaesthetic information, it provides simultaneous information about features of, for instance, a geographical map. It has been shown that such a map has potential for teaching geographical facts to visually impaired readers, but also that revisions can be expected to improve its functioning (Jansson and Pedersen 2005; Jansson *et al.* 2006). A general problem is the use of a tactile mouse without visual feedback, which means that there are problems to be informed about the location and motions of the cursor on the screen. For instance, moving the tactile mouse horizontally with its body oriented non-perpendicularly to the direction of the motion path results in an oblique motion path of the cursor on the screen. There may also be problems to know the location when the mouse is lifted and replaced in a new position, as well as differences in final location of the cursor appearing when the mouse has been moved at different speeds. Anyhow, there have been efforts to investigate the usefulness of haptic mice for visually impaired people in several contexts, such as reading graphs presented both tactually and auditorily (Wall and Brewster 2006), giving directional information during navigation in a virtual environment (Pietrzak *et al.* 2006) and teaching science (Wies *et al.* 2001).

4.7 Haptic Displays

For vision and audition we have for a long time had displays, such as computer screens and loudspeakers that are highly capable of both mirroring reality and producing unseen and unheard material. Not much more than a decade ago, a development of corresponding devices for the sense of haptics began. They are robotic devices allowing both manipulation and information about virtual objects and scenes *via* a computer. Some examples that have been used in visually impaired people's contexts will be discussed here.

The term haptic display has in practice mainly been used for advanced robotic gadgets for haptic exploration of virtual objects, but other devices, such as tactile pictures and the tactile mouse, could be said to belong to this category, but traditionally they are not classified in this way. So far, more than 30 different haptic displays have been constructed (GALLERY, undated). For general overviews, see Burdea (1996), Burdea and Coiffet (2003) and McLaughlin *et al.* (2002).

Most devices developed so far provide a possibility to get force feedback by a "collision" between the endpoint of a stylus, a thimble or suchlike things at the device, held by a user's hand, and the surface of a virtual object. In addition to the shape, surface properties such as hardness/softness, texture and friction can be rendered for haptic exploration. A beginning to render heat properties of the surface has also been made (Monkman and Taylor 1993), even if the rendering of this property is still not well developed.

The most interesting property of the haptic displays as a potential aid for the visually impaired is that they make available direct haptic information about the 3D aspects of an object or a scene, information that is difficult to obtain haptically from 2D depictions of 3D objects and scenes (cf. above). A basic problem for the usefulness of the haptic displays is if this information is sufficient and has a form making it useful. For visually impaired people it is especially important to consider its usefulness for haptics alone, without being combined with vision, and also the possibilities of enhancing haptics with auditory information.

4.7.1 Information Available *via* a Haptic Display

It was stated above that haptics has enormous potentials when functioning in natural contexts. However, these potentials are only partly utilized by haptic displays developed so far. The most important constraints concern *number and size of contact surfaces*. When the bare hand is used naturally, there are several contact surfaces and each of them has an extension of roughly a finger pad. In present day haptic displays the number of contacts is quite low, in most cases just one. The contact surface is also, except in a few experimental devices, only a tiny point. These are drastic differences from natural haptics with important effects for the efficiency of haptic displays.

When there is only one contact area, no simultaneous information from several contact areas can, of course, be obtained; only successive information is available. Using one finger may be natural in some contexts. Symmons and Richardson (2000) found that the use of one finger was very common when exploring a 2D tactile picture. However, the situation is probably different when a 3D object is explored. The use of only one finger means, among other things, that the number of available exploratory procedures decreases. For instance, enclosure, an important exploratory procedure for perceiving global shape where several fingers are grasping an object, cannot be used. Klatzky *et al.* (1993) found a deterioration of performance between five fingers and one finger, and Lederman and Klatzky (2004a) investigated other constraints and provided a theoretical context for the effects of them. Jansson and Monaci (2004) studied the constraints in number of bare fingers to one finger, two, three and five fingers and got related results (Figure 4.6, bare fingers) with the difference between one finger and two fingers being the largest one. However, that is not the whole story.

The constraint to a point-like contact area was not the case in the Jansson and Monaci (2004) study, as pads of the fingers were involved. Lederman and Klatzky (1999) demonstrated the importance of spatially distributed fingertip forces by simulating the restricted information condition at a contact area by rigid sheaths put on the fingertips. Most of the perceptual measures, including tactile spatial acuity, perception of 2D edge detection, showed substantial impairment. If the fingertips are equipped with such rigid sheaths during identification of objects (Figure 4.6, fingers with sheaths), the results are also drastically different (Jansson and Monaci 2006). The difference between one finger and two fingers is still there concerning percent errors, but it is reduced. The large differences are between bare fingers and fingers with sheaths. Amount of information at the contact areas is thus of central importance, more important than number of contacts. The experimental situations in these experiments simulate the situation in presently available potential haptic displays, and they suggest that the effects

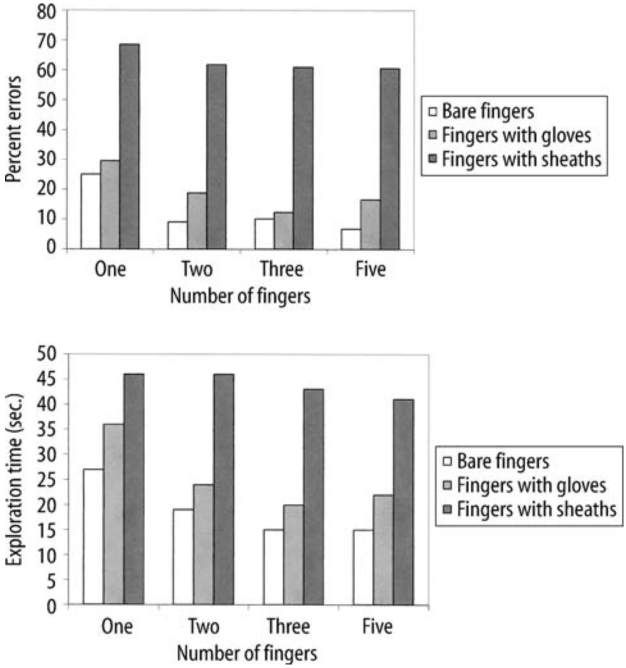


Figure 4.6. Proportion of errors and exploration time as functions of number of fingers and spatially distributed information (Jansson and Monaci 2005)

of the reduced information available *via* haptic displays are very large, especially concerning spatially distributed fingertip information. A lesson for the developers of haptic displays is that not much is won, at least not concerning identification of virtual objects, by increasing the number of contacts, if each of them does not contain spatially distributed information. A study using up to three contact points of a haptic device got the same result of no improvement with number of points (Frisoli *et al.* 2004).

4.7.2 What Information Can Be Obtained with the Reduced Information?

The improvements suggested above can hardly be expected to occur within a near future because of the technical complexity of providing useful extended information at the contact areas. The present display types without the spatially distributed information will probably be dominating for several years. However, there are potentials also with the constrained information, and some relevant for applications to visually impaired people will be discussed below.

4.7.2.1 Identification of Shape

With information reduced to one contact point it is possible to identify simple regular shapes with the help of only haptic information. This was the result of an

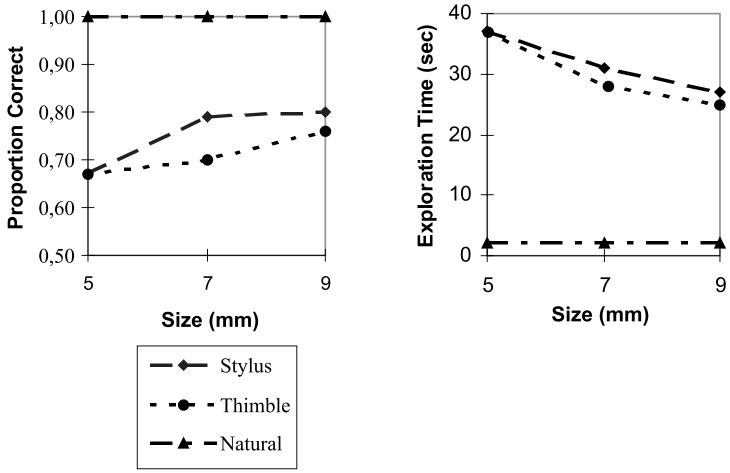


Figure 4.7. Proportion correct and exploration time for identification of simple shapes in three sizes during three conditions (Jansson and Billberger 1999)

experimental study (Jansson *et al.* 1999) using a display allowing exploration of virtual objects *via* a hand-held stylus or a fingertip in a thimble attached to a robot arm, a PHANToM 1.5 (see, PHA, undated). However, the efficiency was much less than that with bare fingers (Figure 4.7). The proportion of correct identifications was perfect in the natural conditions, but only around 0.7 during the two haptic display conditions. The exploration time was much longer during these conditions (around 30 s) than in the natural conditions (about 2 s). It can also be noted that the two haptic display conditions had very similar results. If the participants used the stylus or the thimble version of the PHANToM did not matter. This suggests that the important property when using such a display is the constraint to one contact point, not the optional way of handling it.

Both proportion correct and exploration time in identification of object shapes improve with the size of the object, at least up to 100 mm (Jansson 2000a), but especially exploration time can be expected to increase at larger object shapes. The performance deteriorates when the shapes get more complex (Jansson and Larsson 2002).

4.7.2.2 Judgement of Texture

Textures are important properties for identification of objects. Hollins *et al.* (1993) made multidimensional scaling of textures and found two main dimensions, soft-hard and smooth-rough, and one weaker, sticky-slippery. In contrast with shape, virtual texture seems potentially easy to judge by haptics. In an experiment where judgements of the roughness of real and virtual sandpapers were both explored with a stylus, the judgements were very similar (Jansson *et al.* 1999). Heller (1989) found that textures can be judged by visually impaired people about as well as by sighted people, which suggests that vision is not necessary for texture perception.

Several aspects of perceiving textures *via* a probe were investigated by Lederman *et al.* (1999) and Klatzky *et al.* (2003), and an analysis of many properties of texture perception can be found in Lederman and Klatzky (2004b). The problems of rendering realistic haptic textures were discussed by Choi and Tan (2004).

In spite of the problems in providing realistic material properties to the virtual objects, it is important to make efforts to make such properties easily identifiable, as they are salient for recognizing the objects easily (Klatzky *et al.* 1993; Klatzky and Lederman, in press). An inexpensive way to render some textures, at least roughness, is to use a force-feedback mouse (Klatzky and Lederman 2006; Lederman *et al.* 2006).

4.7.2.3 Effects of Surface Properties on the Perception of Object Form

Virtual objects can be rendered with arbitrary combinations of form and surface properties. Does it matter for the perception of shape what the surface properties are? It can be expected that the form of an object with a soft surface would be more difficult to judge than a hard one because of the risk of distorting the form during exploration. Further, a surface with friction and/or texture may also be more difficult to judge than a surface without these properties, as these surface properties may disturb the exploration. One experiment demonstrated such a result for friction (Christou and Wing 2001). Another experiment suggested that only texture, but not friction, had a significant deteriorating effect (Jansson and Pieraccioli 2004). Concerning softness informal observations in the latter study indicated that the participants adapted their exploratory behaviour to the degree of softness and decreased their pressure on the surface accordingly, which suggests that softness does not deteriorate shape judgments. Further studies should be made to investigate if the disturbing effects for texture can be generalized to other kinds of textures.

4.7.2.4 Training in Exploratory Procedures

One hypothesis why virtual textures can be more successfully perceived than the shape of virtual objects is that the exploratory procedure for judging texture is much simpler than for judging shape. When judging texture it is sufficient just to make any movement over the surface, while it is much more complicated when shape has to be judged. It is reasonable to assume that practice in judging shape should improve performance, which has also been proven experimentally (Jansson and Ivås 2001). Already a few days of 1-h practice sessions demonstrated large improvements for a majority of the participants; one case is shown in Figure 4.8. It should also be noted that a minority of the participants did not show any improvement, which was interpreted to depend on insufficient motivation. An important practical conclusion of this result is that there is a great risk of underestimating the usefulness of a haptic display, if the participants have not sufficient practice in exploration with the specific display.

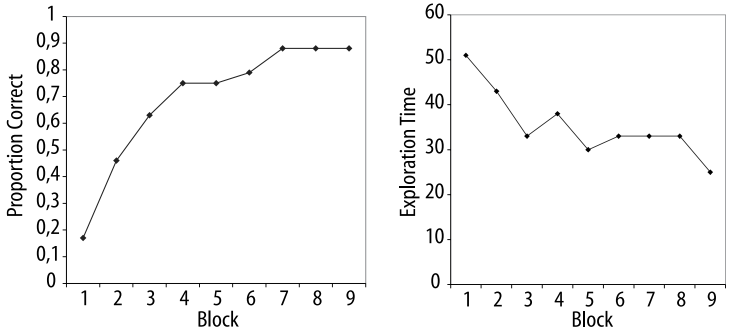


Figure 4.8. Practice results in proportion correct and exploration time (s) for a very successful participant in an experiment with identification of four forms in six different sizes. The practice consisted of three blocks of stimuli during each of three days (Can the Efficiency of a Haptic Display be Increased by Short-time Practice in Exploration?, Jansson G., Ivås A., in Haptic Human–Computer Interaction, Brewster S., Murray-Smith R., 2001, © Springer-Verlag. Used with permission.)

4.7.2.5 Co-location of Visual and Haptic Space

In natural environments objects are usually perceived to occupy the same location visually and tactually. In virtual worlds they are often not co-located, that is, the visual presentation is made on a screen and the tactual presentation in another location. This may be a drawback not only for sighted people, but also for visually impaired people with remains of vision. It is reasonable to expect advantages with co-location, as that is the natural condition. Some such effects were shown for a targeting task (Wall *et al.* 2002) and for perception of form (Jansson and Öström 2004). Under stereo conditions in the latter study there was a significant effect in judging the depth dimension of a distorted object. Informally it was also found that performing tasks such as finding knobs and regaining contact with lost virtual objects were facilitated under co-location conditions.

4.7.3 Haptic Displays as Aids for the Visually Impaired

There are many potential applications of haptic displays, for instance, in medical training and for computer games (for an overview, see Burdea and Coiffet 2003). Here their potentials as aids for the visually impaired will be discussed. Users with severe visual impairment have to work without the extra information obtained from vision, which makes the task more difficult, especially concerning getting an overview and not loosing contact with an object explored, as well as finding interesting parts of the object to touch. In spite of this, haptic displays may be useful for visually impaired people, as it is the only way in some contexts for getting direct information (Brewster and Murray-Smith 2001; Jansson 1999, 2001; Sjöström 2002; Yu *et al.* 2001).

One context where haptic displays may be useful for visually impaired people is to experience art at museums. It is usually not allowed to touch statues at museums, for instance. The EU project Pure-Form (PFM, undated) aimed to develop a haptic display for the exploration of virtual copies of statues (Bergamasco *et al.* 2001; Bergamasco and Prisco 1998; Frisoli *et al.* 2002). The Pure-Form display as a technical device is shown in Figure 4.9 and used in a museum context in Figure 4.10. This display is intended for a general group of users, but it has, of course, a special interest for visually impaired people as a unique possibility of a direct experience of statues on their own (Jansson *et al.* 2003). A prototype has been evaluated at four European museums (Frisoli *et al.* 2005). Sighted participants found in general the haptic experience amusing and instructive, and so did the visually impaired participants, but to a somewhat lower degree. The usefulness for

Figure 4.9. The Pure-Form display carried by a user. Image copyright Antonio Frisoli, Massimo Bergamasco, PERCRO, Scuola Superiore Sant'Anna, Pisa, Italy and used by permission

Figure 4.10. The Pure-Form display used together with a visual display in a museum. Image copyright Antonio Frisoli, Massimo Bergamasco, PERCRO, Scuola Superiore Sant'Anna, Pisa, Italy and used by permission

the visually impaired was not as clear as expected. The reason is probably that the haptic display, in spite of the increased 3D information it provided, did not solve the general problem of haptic perception without vision, namely that of obtaining an overview of an object or scene. In order to be maximally useful for users without vision some help with the overview is needed, for instance, verbal introductory explanation and/or guidance for exploration of the object or scene. A suggestion to help the visually impaired to be oriented in a complex virtual context is to provide an external memory aid (Wall and Brewster 2004).

In another EU project, known as GRAB (GRAB, undated), a new haptic and audio virtual environment was developed. In the GRAB haptic display, the user puts two fingers in thimbles attached to two arms providing 3D force feedback. For the GRAB system three applications were developed and evaluated: a game, a chart data explorer and a city-map explorer. The evaluations indicated its usefulness for visually impaired people concerning these applications, as well as they gave suggestions for improvements. An example is the evaluation of the maps application (GRAB 2004).

A general problem concerning technically complicated devices such as these is that they may have shown their potential in laboratory contexts, but the usefulness in real situations may be more problematic. The intended users may find them difficult to use, possibly because they are not sufficiently adapted to the functioning of haptics (Hatwell 2006). To make good psychophysical studies is difficult because of the need to choose good methods, find exact definitions of proximal stimuli, and handle new hardware. Therefore close cooperation between several sciences, including psychology and engineering, is important (Tan 2006). Guidelines for the tactile and haptic interaction between computer users and relevant devices is under development by ISO (International Standards Organization) (van Erp *et al.* 2006).

4.8 Chapter Summary

Historically the sense of touch has been used extensively to generate information for the visually impaired person. This chapter surveyed the underlying fundamental principles and the perceptual capabilities of haptics achievable with the human hand.

After the presentation and explanation of these haptic fundamentals, the chapter proceeds to investigate how haptics can be used and enhanced through training or with the aid of specialist tools. A central section of the chapter concentrated on *low-tech* haptic applications; some, like the long cane and the guide dog, were for mobility whilst others, like Braille and embossed pictures, were for information from text, as well as embossed graphics. An important feature of this section was to identify lessons from the use of low-tech haptic assistive technology.

Subsequent sections in the chapter examined the more technologically advanced applications of haptic science. Of particular importance were the technologies for haptic computer interfaces and for haptic displays. A project to provide haptic access to museum statues and artefacts for visually impaired people was one demonstrable outcome of this advanced work.

4.9 Concluding Remarks

Haptics is a very capable sense whose potentials have not been generally recognized. However, it does not have the capacity of vision to get an immediate overview and to cover space far away. Its strength is in near-space where the bare hands in real environments can very competently perform many tasks, including, for instance identification of objects and judging surface properties. This capacity has been utilized in several low-tech aids for the visually impaired. The success of many of them is a memento, demonstrating that the important aspect of devices for the blind is that they are adapted to the functioning of the replacing sense. More technically advanced aids have been developed but with more moderate success. They make in many cases information available that would be very helpful if it were sufficiently adapted to the senses in question. It is a great challenge to develop aids that present really useful information to the users and make possible a well functioning interaction between user and aid.

Acknowledgement. The author's research reported here was funded by the Swedish Council for Research in Humanities and Social Sciences, the Bank of Sweden Tercentenary Fund, the Swedish Transport and Communication Research Board and the EU projects TIDE-BP-1148 – MOBIC, IST-2000-29580-PURE-FORM, IST-2003-511592 STP MICOLE and IST-2004-002114-ENACTIVE.

Questions

Q.1 What are the most important characteristics of the haptic sense?
Q.2 What are the differences between the use of haptics in natural contexts and with aids for the visually impaired?
Q.3 Discuss the reasons for the relatively larger problems using technically advanced aids, as compared with those using low-tech aids?
Q.4 How can graphics be made accessible to those who cannot see?
Q.5 How can auditory information contribute to the usefulness of haptic information when vision is not available?
Q.6 In what ways can haptic displays be useful as aids for visually impaired people, and what remains to be done to improve their usefulness?

Projects

P.1 There are several websites discussing accessibility problems for visually impaired people where many different ways are described. Sometimes there are suggestions including haptics, sometimes haptics is not mentioned. Find out what haptic options are mentioned and look for cases where such options are not mentioned but where you can find out suggestions. Here are some major addresses: American Printing House for the Blind (APH, undated), American Foundation for the Blind (AFB, undated), Royal National Institute of the Blind (RNIB, undated) and Trace Research & Development Center (TRACE, undated).

P.2 Discuss if there are any differences in usefulness for visually impaired users of the different haptic displays (GALLERY, undated). The usefulness may vary for different kinds of tasks, such as getting web accessibility, reading 2D graphics, and identifying virtual 3D objects and their properties (shape, texture, hardness, weight).

P.3 As discussed in the text, getting spatially distributed information to the contact areas of the haptic displays is the most important way of improving their functioning for identification of virtual objects. Search on the Web for suitable solutions for this requirement. Observe that even if equidistant pins may be a solution in the right direction, technology that would allow varying distances between sensation points probably is a still better solution, as stimulus variables such as texture gradients important for depth perception would be possible.

References

AFB, undated, American Foundation for the Blind, http://www.afb.org

APH, undated, American Printing House for the Blind, http://www.aph.org

Austin, T., and Sleight, R., 1952, Accuracy or tactual discrimination of letters, numerals, and geometric form, J. Experimental Psychology, Vol. 43, pp. 239–247

Bach-y-Rita, P., 1972, Brain Mechanisms in Sensory Substitution, Academic Press, New York

Bach-y-Rita, P., and Kercel, S.W., 2003, Sensory substitution and the human–machine interface, Trends in Cognitive Sciences, Vol. 7, pp. 541–546

Bach-y-Rita, P., Kaczmarek, K., Tyler, M., and Garcia-Lara, J., 1998, Form perception with a 49-point electrotactile stimulus array on the tongue, J. Rehabilitation Research and Development, Vol. 35, pp. 427–431

Bach-y-Rita, P., Kaczmarek, K.A., and Tyler, M.E., 2003, A tongue-based tactile display for portrayal of environmental characteristics, In J. Hettinger and M.W. Haas (Eds.), Virtual and adaptive environments: Applications, implications, and human performance, Erlbaum, Mahwah, NJ

Ballesteros Jiménez, S., and Heller, M.A. (Eds.), 2004, Touch, blindness and neuroscience, UNED Press, Madrid, Spain

Barth, J.L., 1982, The development and evaluation of a tactile graphics kit, J. Visual Impairment and Blindness, Vol. 76, pp. 269–273

Barth, J.L., and Foulke, E., 1979, Preview: A neglected variable in orientation and mobility, J. Visual Impairment and Blindness, Vol. 73, pp. 41–48

Bentzen, B.L., 1997, Orientation aids, In B.B. Blash, W.R. Wiener and R.L. Welsh (Eds.), Foundations of orientation and mobility, 2nd Edn., pp. 284–316. AFB Press, New York

Bergamasco, M., and Prisco, G., 1998, Design of an anthropomorphic haptic interface for the human arm, In Y. Shirai and S. Hircse (Eds,), Robotic Research, the Eight International Symposium, pp. 278–289, Springer, London, UK

Bergamasco, M., Avizzano, C., Di Petri, G., Barbagli, F., and Frisoli, A., 2001, The museum of pure form: system architecture, In Procs of 10th IEEE International Workshop on Robot and Human Interactive Communication, pp. 112–117, IEEE Press, Piscataway, NY

Blash, B.B., Wiener W.R., and Welsh R.L. (Eds.), 1997, Foundations of orientation and mobility, 2nd Edn., AFB Press, New York

Bledsoe, C.W., 1997, Originators of orientation and mobility training, In B.B. Blash, W.R. Wiener and R.L. Welsh (Eds.), Foundations of orientation and mobility, 2nd. Edn., pp. 580–623. AFB Press, New York

Bliss, J.C., 1978, Reading machines for the blind, In G. Gordon (Ed.), Active touch: The mechanism of recognition of objects by manipulation, A multidisciplinary approach, Pergamon Press, Oxford, UK

Brewster, S., and Murray-Smith, R. (Eds.), 2001, Haptic human-computer interaction, Springer, Berlin, Germany

Burdea, G., 1996, Force and touch feedback for virtual reality, Wiley, New York, USA

Burdea, G., and Coiffet, P., 2003, Virtual Reality Technology (2nd Edn. with CD-ROM), Wiley, New York

Burger, D., Mazurier, C., Cesarano, S., and Sagot, J., 1993, In D. Burger and J.-C. Sperandio, Non-visual human-computer interactions, Prospects for the visually handicapped, Colloque INSERM, Vol. 228, pp. 97–114, John Libbey Eurotext, Montrouge, France

Choi, S., and Tan, H.Z., 2004, Toward realistic haptic rendering of surface textures, IEEE Computer Graphics and Applications, Vol. 24, pp. 40–47

Cholewiak, R., and Collins, A., 1991, Sensory and physiological basis of touch, In M. A. Heller and W. Schiff (Eds.), The psychology of touch, Erlbaum, Hillsdale, NJ, USA

Chomsky, C., 1986, Analytic study of the Tadoma method: Language abilities of three deaf-blind subjects, J. Speech and Hearing Research, Vol. 29, pp. 332–347

Christou, C., and Wing, A., 2001, Friction and curvature judgment, In C. Baber, M. Faint, S. Wall and A.M. Wing (Eds.), Eurohaptics 2001 Conf. Procs. (Educational Technology Research Papers, ETRP 12, ISSN 1463-9394). The University of Birmingham, Birmingham, England, pp. 36–40

Cratty, B.J., 1971, Movement and spatial awareness in blind children and youth, Thomas, Springfield, IL, USA

Cronin, V., 1977, Active and passive touch at three age levels, Developmental Psychology, Vol. 13, pp. 253–256

DMD, undated, DMD 120060, http://www.metec-ag.de/company.html

Edman, P.K., 1992, Tactile graphics, American Foundation for the Blind, New York

Eriksson, Y., 1998, Tactile pictures: Pictorial representations for the blind 1784–1940, Acta Universitatis Gothoburgensis, Gothenburg Studies in Art and Architecture, No. 4.

Farmer, L.W., and Smith D.L., 1997, Adaptive technology, In B.B. Blash, W.R. Wiener and R.L. Welsh (Eds.), Foundations of orientation and mobility (2nd. Edn.), pp. 231–259, AFB Press, New York

Foulke, E., 1991, Braille, In M. A. Heller and W. Schiff (Eds.), The psychology of touch, pp. 219–233, Erlbaum, Hillsdale, NJ

Frisoli, A., Simoncini, F., and Bergamasco, M., 2002, Mechanical design of a haptic interface for the hand, In Procs. 2002 ASME DETC 27th Biennial Mechanisms and Robotics Conference, Montreal, Canada, pp. 25–32, ASME Press, Montreal, Canada

Frisoli, A., Barbagli, F., Wu, S.-L., Ruffaldi, E. Bergamasco, M., and Salisbury, K., 2004, Evaluation of multipoint contact interfaces in haptic perception of shapes, Manuscript

Frisoli, A., Jansson, G., Bergamasco, M., and Loscos, C., 2005, Evaluation of the Pure-Form haptic displays used for exploration of works of art at museums. Procs of Worldhaptics 2005, Pisa, March 18–20, 2005, Available in Procs on CD-ROM

GALLERY, undated, http://haptic.mech.northwestern.edu/intro/gallery

Gesink, J., Guth, D., and Fehr, B., 1996, A new, talking, gyroscopic device for training blind pedestrians to walk a straight path and make accurate turns, In J.M. Tellevik and G.E. Haugum (Eds.), Conf. Procs: International Mobility Conference No. 8, Trondheim and Melhus, Norway, May 10-19, 1996, pp. 108–110, Tambartun National Resource Center for Special Education of the Visually Handicapped, Tambartun, Norway

Gibson, J.J., 1962, Observations on active touch, Psychological Review, Vol. 69, pp. 477–491

Gill, J.M., 1973, Design, production and evaluation of tactual maps for the blind, Unpublished Ph. D. thesis, University of Warwick, Warwick, UK

Gill, J.M., 1982, Production of tangible graphic displays, In W. Schiff and E. Foulke, Tactual perception: A sourcebook, pp. 405–416, Cambridge University Press, Cambridge, UK

GRAB, undated, GRAB IST-2000-26151, http://www.grab-eu.com

GRAB, 2004, Results of the validation of the maps application, GRAB IST-2000-26151, Deliverable 16/3

Guarniero, G., 1974, Experience of tactile vision, Perception, Vol. 3, pp. 101–104

Guarniero, G., 1977, Tactile vision: a personal view. Journal of Visual Impairment and Blindness, Vol. 71, pp. 125–130

GUIB project, 1995, Textual and graphical user interfaces for blind people, Final report, Royal National Institute for the Blind, London, UK

Guth, D.A., and LaDuke, R., 1994, The veering tendency and blind pedestrians: An analysis of the problem and literature review, J. Visual Impairment and Blindness, Vol. 88, pp. 391–400

Guth, D.A., and Rieser, J.J., 1997, Perception and the control of locomotion by blind and visually impaired pedestrians, In B.B. Blash, W.R. Wiener and R.L. Welsh (Eds.), Foundations of orientation and mobility,(2nd. Edn.), pp. 9–38, AFB Press, New York

Guth, D.A., Hill, E., and Rieser, J., 1989, Tests of blind pedestrians' use of traffic sounds for street crossing alignment, J. Visual Impairment and Blindness, Vol. 83, pp. 461–468

Guth, D.A., LaDuke, R., and Gesink, J., 1996, Is feedback about body rotation useful for decreasing veer and increasing the accuracy of turns? Test of a new device, In J.M. Tellevik and G.E. Haugum (Eds.), Conf. Procs: Int. Mobility Conf. No. 8, Trondheim and Melhus, Norway, May 10–19, pp. 111–113, Tambartun National Resource Center for Special Education of the Visually Handicapped, Tambartun, Norway

Hännestrand, B., 1995, Människan, samhället och ledarhunden: studier i ledarhundens historia (Man, society and studies in the history of work with guide dogs), Ph. D. thesis, Uppsala Studies in Economic History, 36 (abstract in English)

Hardwick, A.J., 2004, Rendering of Moon text on simulated tactile diagrams for blind computer users by force-feedback, In S. Ballesteros Jiménez and M.A. Heller (Eds.), Touch, Blindness and Neuroscience, pp. 351–358, UNED Press, Madrid, Spain

Hatwell, Y., 2006, A survey of some contemporary findings on haptic perception, Plenary talk at the Eurohaptics Int. Conf., EH 2006, July 3–6, Paris France, Procs. p. 3 (abstract)

Heller, M.A., 1989, Texture perception in sighted and blind observers, Perception and Psychophysics, Vol. 45, pp. 49–54

Heller, M.A. (Ed.), 2000, Touch, representation and blindness, Oxford University Press, Oxford, UK

Heller, M.A., and Schiff, W. (Eds), 1991, The psychology of touch, Erlbaum, Hillsdale, NJ

Hollins, M., Faldowski, R., Rao, S., and Young, F., 1993, Perceptual dimensions of tactile surface texture: A multidimensional scaling analysis, Perception and Psychophysics, Vol. 54, 697–705

Holmes, E., Michel, R., and Raab, A., 1995, Computerunstützte Erkundung digitaler Karten dursch Sehbehinderte (Computer supported exploration of digital maps for visually impaired people), In W. Laufenberg and J. Lötzsch (Eds.), Taktile Medien. Kolloquium über tastbare Abbildungen für Blinde, pp. 81–87, Deutsche Blindenstudienanstalt e.V., Carl-Strehl-Schule, Marburg, Germany

Holmes, E., Jansson, G., and Jansson, A., 1996, Exploring auditorily enhanced maps for travel in new environments, In D. Burger (Ed.), New technologies in the education of the visually handicapped, pp. 191–196, John Libbey Eurotext Montrouge, France

Holmes, E., Hughes, B., and Jansson, G., 1998, Haptic perception of texture gradients, Perception, Vol. 27, pp. 993–1008

Hughes, B., and Jansson, G., 1994, Texture perception via active touch, Human Movement Science, Vol. 13, pp. 301–333

James, G., and Armstrong, J.D., 1976, Handbook on mobility maps, Mobility Monographs, No 2, Blind Mobility Research Unit, Department of Psychology, University of Nottingham, Nottingham, UK

Jansson, G., 1972, Symbols for tactile maps, In B. Lindquist and N. Trovald (Eds.), European Conf. on Educational Research for the Visually Handicapped (Rep No. 31, pp. 66–77), The Uppsala Institute for Education, Dept. of Education, Uppsala, Sweden

Jansson, G., 1983, Tactile guidance of movement, Int. J. Neuroscience, Vol. 19, pp. 37–46

Jansson, G., 1993, Perception of the amount of fluid in a vessel shaken by hand, In S.S. Valenti and J.B. Pittinger (Eds.), Studies in perception and action II, Posters presented at the VIIth Int. Conf Event Perception and Action. August 8-13, 1993. University of British Columbia, Vancouver, BC, Canada, pp. 263–267, Erlbaum, Hillsdale, NJ

Jansson, G., 1995, Information about direction – a comparison of verbal and vibro-tactile information forms, In E. Holmes, A. Jansson, G. Jansson, V. Johnson and H. Petrie (Eds.), Report on full evaluation of prototype interfaces for the MoBIC Travel Aid. Unpublished report. TIDE project 1148-MoBIC

Jansson, G., 1999, Can a haptic display rendering virtual 3D objects be useful for people with visual impairment?, J. Visual Impairment and Blindness, Vol. 93, pp. 426–429

Jansson, G., 2000a, Basic issues concerning visually impaired people's use of haptic displays, In P. Sharkey, A. Cesarani, L. Pugnatti and A. Rizzo (Eds.), 3rd Int. Conf. Disability, Virtual Reality and

Associated Technologies – Procs, pp. 33–38, 23–25 September, Alghero, Sardinia, Italy, University of Reading, Reading, UK, Also available at http://www.icdvrat.reading.ac.uk/2000/papers/2000_05.pdf

Jansson, G., 2000b, Spatial orientation and mobility of people with vision impairment, In B. Silverstone, M.A., Lang, B.P., Rosenthal and E.E. Faye (Eds.), The Lighthouse handbook on visual impairment and vision rehabilitation, pp. 359–375, Oxford University Press, New York

Jansson, G., 2001, The potential usefulness of high-tech aids for visually impaired seniors, In H.-W. Wahl and H.-E. Schulze (Eds.), On the special needs of blind and low vision seniors, pp. 231–238, IOS Press, Amsterdam

Jansson, G., 2003, Tactile maps – overview of research and development, In Y. Eriksson, G. Jansson and M. Strucel (Eds.), Tactile maps, Guidance in map production, pp. 45–78, The Swedish Braille Authority, Stockholm, Sweden

Jansson, G., and Billberger, K., 1999, The PHANToM used without visual guidance, In Proceedings of the First PHANToM Users Research Symposium (PURS99), May 21–22, 1999, Deutsches Krebsgorschungszentrum, Heidelberg, Germany

Jansson, G., and Brabyn, L., 1981, Tactually guided batting, Uppsala Psychological Reports, No. 304, Department of Psychology Uppsala University, Uppsala, Sweden

Jansson, G., and Holmes, E., 2003, Can we read depth in tactile pictures? Potentials suggested by research in tactile perception, In E. Axel and N. Levant (Eds.), Art beyond vision: A resource guide to art, creativity, and visual impairment, pp. 146–156, Art Education for the Blind and American Foundation for the Blind, New York

Jansson, G., and Ivås, A., 2001, Can the efficiency of a haptic display be increased by short-time practice in exploration?, In S. Brewster and R. Murray-Smith, Haptic Human-Computer Interaction, pp. 85–91, Springer, Heidelberg, Germany

Jansson, G., and Larsson, K., 2002, Identification of haptic virtual objects with different degrees of complexity, In S.A. Wall, B. Riedel, A Crossan and M.R. McGee (Eds.), Eurohaptics 2002, Conf. Procs, pp. 57–60, Edinburgh University, Edinburgh, UK

Jansson, G., and Monaci, L., 2004, Haptic identification of objects with different numbers of fingers, In S. Ballesteros Jiménez and M.A. Heller (Eds.), Touch, Blindness and Neuroscience, pp. 203–213, UNED Press, Madrid, Spain

Jansson, G., and Monaci, L., 2005, Improving haptic displays: Providing differentiated information at the contact areas is more important than increasing the number of areas. Poster at Worldhaptics 05, Pisa, March 18–20 2005. Available in Proceedings on CD-ROM

Jansson, G., and Monaci, L., 2006, Identification of real objects under conditions similar to those in haptic displays: Providing spatially distributed information at the contact areas is more important than increasing the number of areas, Virtual Reality, Vol. 9, 243–249

Jansson, G., and Öström, M., 2004, The effects of co-location of visual and haptic space on judgements of forms, In M. Buss and M. Fritschi (Eds.), Procs 4th Int. Conf Eurohaptics 2004, pp. 516–519, Technische Universität München, München, Germany

Jansson, G., and Pedersen, P., 2005, Obtaining geographical information from a virtual map with a haptic mouse, International Cartographic Conference (Theme "Maps for Blind and Visually Impaired"), A Coruña, Spain, July 9–16, Available on conference CD-ROM

Jansson, G., and Pieraccioli, C., 2004, Effects of surface properties on the perception of the form of virtual objects, In M. Buss and M. Fritschi (Eds.), Procs 4th Int. Conf. Eurohaptics 2004, pp. 211–216, Technische Universität München, München, Germany

Jansson, G., Billberger, K., Petrie, H., Colwell, C., Kornbrot, D., Fänger, J., König, H., Hardwick, A., and Furner, S., 1999, Haptic virtual environments for blind people: Exploratory experiments with two devices, Int. J. Virtual Reality, Vol. 4, pp. 10–20

Jansson, G., Bergamasco, M., and Frisoli, A., 2003, A New Option for the Visually Impaired to Experience 3D Art at Museums: Manual Exploration of Virtual Copies, Visual Impairment Research, Vol. 5, pp. 1–12

Jansson, G., Juhasz, I., and Cammilton, A., 2006, Reading virtual maps with a haptic mouse: Effects of some modifications of the tactile and audio-tactile information, British J. of Visual Impairment, Vol. 24, 60–66

Jansson, G., Juslin, P., and Poom, L., 2006, Liquid-specific properties can be utilized for haptic perception of amount of liquid in a vessel put in motion, Perception, Vol. 35, pp. 1421–1432

Johnson, K., 2002, Neural basis of haptic perception, In H. Pashler and S. Yantis, Stevens' Handbook of Experimental Psychology, Vol. 1. Sensation and Perception, 3rd Edn., pp. 537–583, Wiley, New York
Kaczmarek, K.A., and Bach-y-Rita, P., 1995, Tactile displays, In W. Barfied and T. Furness III (Eds.), Virtual environments and advanced interface design, pp. 349–414, Oxford University Press, New York
Katz, D., 1989, The world of touch, (Original work published 1925; Translated by L.E. Kreuger), Erlbaum, Hillsdale, NJ
Kelly, T., and Schwartz, L., 1999, "Talking kiosk" for the visually impaired unveiled at Penn station, Pamphlet
Kennedy, J.M., 1993, Drawing and the blind: Pictures to touch, Yale University Press, New Haven, CT
Kennedy, J.M., Gabias, P., and Nicholls, A., 1991, In M. A. Heller and W. Schiff (Eds.), The psychology of touch, pp. 263–299, Erlbaum, Hillsdale, NJ
Klatzky, R.L., and Lederman, S.J., 1995, Identifying objects from a haptic glance, Perception and Psychophysics, Vol. 57, pp. 1111–1123
Klatzky, R.L., and Lederman, S.J., 2006, The perceived roughness of resistive virtual textures: I. Rendering by a force-feedback mouse, ACM Transactions on Applied Perception, Vol. 3, pp. 1–14
Klatzky, R.L., and Lederman, S.J., 2007, Object recognition by touch, In J. Rieser, D. Ashmeed, F. Ebner and A. Com (Eds.), Blindness and brain plasticity in navigation and object perception, Mahwah, Erlbaum
Klatzky, R.L., Lederman, S.J., and Metzger, V.A., 1985, Identifying objects by touch: An "expert system", Perception and Psychophysics, Vol. 37, pp. 299–302
Klatzky, R.L., Loomis, J.M., Lederman, S.J., Wake, H., and Fujita, N., 1993, Haptic identification of objects and their depictions, Perception and Psychophysics, Vol. 54, pp. 170–178
Klatzky, R.L., Lederman, S.J., Hamilton, C., Grindley, M., and Swendsen, R.H., 2003, Feeling textures through a probe: Effects of probe and surface geometry and exploratory factors, Perception and Psychophysics, Vol. 65, pp. 613–631
Lederman, S., 1981, The perception of surface roughness by active and passive touch, Bull. Psychonomic Society, Vol. 18, pp. 253–255
Lederman, S.J., and Campbell, 1982, Tangible graphs for the blind, Human factors, Vol. 24, pp. 85–100
Lederman, S.J., and Kinch, D.H., 1979, Texture in tactual maps and graphics for the visually handicapped, Visual impairment and Blindness, Vol. 73, pp. 217–227
Lederman, S.J., and Klatzky, R.L., 1987, Hand movements: A window into haptic object recognition, Cognitive Psychology, Vol. 19, pp. 342–368
Lederman, S.J., and Klatzky, R.L., 1999, Sensing and displaying spatially distributed fingertip forces in haptic interfaces for teleoperators and virtual environment systems, Presence, Vol. 8, pp. 86–103
Lederman, S.J., and Klatzky, R.L., 2004a, Haptic identification of common objects: Effects of constraining the manual exploration process. Perception and Psychophysics, Vol. 66, pp. 618–628
Lederman, S.J., and Klatzky, R.L., 2004b, Multisensory texture perception, In G.A. Calvert, C. Spence and B.E. Stein (Eds.), Handbook of multisensory processes, pp. 107–122, MIT Press, Cambridge, MA
Lederman, S.J., Klatzky, R.L., Hamilton, C.L., and Ramsey, G.I., 1999, Perceiving roughness via a rigid probe: Psychophysical effects of exploration speed and mode of touch, Haptics-e, Vol. 1, pp. 1–20
Lederman, S.J., Klatzky, R.L., Tong, C., and Hamilton, C, 2006, The perceived roughness of resistive virtual textures: II. Effects of varying viscosity with a force-feedback device, ACM Transactions on Applied Perception, Vol. 3, pp.15–30
Leotta, D.F., Rabinowitz, W.M., Reed, C.M., and Durlach, N.I., 1988, Preliminary results of speech-reception tests obtained with the synthetic Tadoma system, J. Rehabilitation Research and Development and Development, Vol. 25, pp. 45–52
Lötzsch, J., 1995, Von audio-taktilen Grafiken zu interaktiven 3D-Modellen (From audio-tactile Graphics to interactive 3D models), In W. Laufenberg and J. Lötzsch (Eds.), Taktile Medien. Kolloquium über tastbare Abbildungen für Blinde, pp. 130–136, Deutsche Blindenstudienanstalt e.V., Carl-Strehl-Schule, Marburg, Germany
Loo, C.K.C., Hall, L.A., McCloskey, D.I., and Rowe, M.J., 1983, Proprioceptive contributions to tactile identification of figures: Dependence on figure size, Behavioral Brain Research, Vol. 7, pp. 383–386
Magee, L.E., and Kennedy, J.M., 1980, Exploring pictures tactually, Nature, Vol. 283, pp. 287–288

McLaughlin, M.L., Hespanha, J.P., and Sukhatme, G.S., 2002, Touch in virtual environments: Haptics and the design of interactive systems, Prentice Hall, Upper Saddle River, NJ

Metria, 2003, Utveckling av Tactile Mapper och Tactile GIS (Development of Tactile Mapper and Tactile GIS), Metria, Kiruna, Sweden

Michel, R., 1999, Interaktiver Layoutentwurf für individuelle taktile Karten (Interactive layout plan for individual tactile maps), Ph.D. thesis, Der Fakultät für Informatik, Otto-von-Guericke Universität, Magdeburg, Germany

Millar, S., 1997, Reading by touch, Routledge, London, UK

Monkman, G., and Taylor, P., 1993, Thermal tactile sensing, IEEE Trans. Robotics and Automatation, Vol. 9, pp. 313–318

Öhngren, G., 1992, Touching voices: Components of direct tactually supported speechreading, Acta Universitatis Upsaliensis, Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences, 32, Uppsala University, Uppsala, Sweden

Palacz, O., and Kurcz, E., 1978, Przydatnos'c' zmodyfikowanego elektroftalmu EWL-300 wg Starkiewicza dla niewidomych (The usefulness of modified Electrophthalm designed be Starkiewicz for the blind), Klinika Oczna, Vol. 48, pp. 61–63

PAP, undated, http://www.papenmeier.de/reha/rehae.htm

Parkes, D., 1988, "Nomad": An audio-tactile tool for the acquisition, use and management of spatially distributed information by visually impaired people, In A.F. Tatham and A.G. Dodds (Eds.), Procs 2nd Int. Symp. on Maps and Graphics for Visually Impaired People, pp. 24–29, London, UK

Pawlik, D.T.V., van Buskirk, C.P., Killebrew, J.H., Hsiao, S.S., and Johnson, K.O., 1998, Control and pattern specification for a high density tactile array, Procs. ASME Dynamic Systems and Control Division, Vol. DSC-64, American Society of Mechanical Engineers (http://www.asme.org)

PFM,-, IST-2000-29580-PURE-FORM, http://www.pureform.org

PHA, undated, http://www.sensable.com

Pietrzak, T., Pecci, I., and Martin, B., 2006, Static and dynamic tactile cues experiments with VTPlayer nouse, Procs. Eurohaptics Int. Conf., EH 2006, July 3–6, 2006, Paris, France, pp. 63–69

Rabinozitz, W.M., Henderson, D.R., Reed, C.M., Delhorne, L.A., and Durlach, N.I., 1990, Continuing evaluation of a synthetic Tadoma system, J. Acoustical Society of America, Vol. 87, pp. 88

Reed, C.M., Rabinowitz, W.M., Durlach, N.I., Braida, L.D., Conway-Fithian, S., and Schultz, M.C., 1985, Research on the Tadoma method of speech communication, J. Acoustical Society of America, Vol. 77, pp. 247–257

Reed, C.M., Durlach, N.I., and Delhorne, L.A., 1992, Natural methods of tactual communication, In I. A. Summers (Ed.), Tactile aids for the hearing impaired, pp. 218–230, Whurr Publishers, London, UK

Richardson, B.L., Symmons, M.A., and Wuillemin, D.B., 2004, The relative importance of cutaneous and kinesthetic cues in raised line drawings identification, In S. Ballesteros Jiménez and M. A. Heller (Eds.), Touch, blindness, and neuroscience, pp. 247–250, UNED Press, Madrid, Spain

RNIB, undated, Royal National Institute of the Blind, http://www.rnib.org

Runeson, S., and Frykholm, G., 1981, Visual perception of lifted weight, J. Experimental Psychology: Human Perception and Performance, Vol. 7, pp. 733–740

Schenkman, B., 1985, Human echolocation: The detection of objects by the blind, Acta Universitatis Upsaliensis, Abstracts of Uppsala Dissertations from the Faculty of Social Sciences 36, Uppsala University, Uppsala, Sweden

Schenkman, B., and Jansson, G., 1986, The detection and localization of objects by the blind with the aid of long cane tapping sounds, Human Factors, Vol. 28, pp. 607–618

Schiff, W., and Foulke, E., 1982, Tactual perception: A sourcebook, Cambridge University Press, Cambridge, UK

Schiff, W., and Isikow, H., 1966, Stimulus redundancy in the tactile perception of histograms, Int. J. for the Education of the Blind, Vol. 16, pp. 1–11

Schultz, M.C., Norton, S.J., Conway-Fithian, S., and Reed, C.M., 1984, A survey of the use of the Tadoma method in the United States and Canada, Volta Review, Vol. 68, pp. 733–737

Shinohara, M., Shimizu, Y., and Mochizuki, A., 1998, Three-dimensional tactile display for the blind, IEEE Trans. Rehabilitation Engineering, Vol. 6, pp. 249–256

Sjöström, C., 2002, Non-visual hapic interaction design: Guidelines and applications, Ph.D. thesis, Number 2:2002, Division of Rehabilitation Engineering Research, Department of Design Sciences (CERTEC), Lund Institute of Technology, Lund, Sweden

Starkiewicz, W., and Kuliszewski, Y., 1963, Active energy radiating system: the 80-channel elektroftalm, In Procs International Congress on Technology and Blindness, American Foundation for the Blind, New York

Stevens, J.C., 1992, Aging and spatial acuity of touch, J. Gerontology, Vol. 47, pp. 35–40

Stevens, J.C., Folke, E., and Patterson, M., 1996, Tactile acuity, ageing, and Braille reading in long-term blindness, J. Experimental Psychology: Applied, Vol. 2, pp. 91–106

Summers, I.R. (Ed.), 1992, Tactile aids for the hearing impaired, Whurr Publishers, London, UK

Symmons, M. and Richardson, B., 2000, Raised line drawings are spontaneously explored with a single finger, Perception, Vol. 29, pp. 621–626

Symmons, A.A., Richardson, B.L., and Wuillemin, D.B., 2004, Active versus passive touch: Superiority depends more on the task than the mode, In S. Ballesteros Jiménez and M.A. Heller (Eds.), Touch, blindness, and neuroscience, pp. 179–185, UNED Press, Madrid, Spain

Tan, H.Z., 2006, The role of psychophysics in haptic research: An engineer's perspective, Plenary talk at Eurohaptics Int. Conf., EH 2006, July 3–6, Paris, France, Procs. p. 5 (abstract)

Thayer, S., 1982, Social touching, In W. Schiff and E. Foulke, Tactual perception: A sourcebook, pp. 263–304, Cambridge University Press, Cambridge, UK

TRACE, undated, TRACE R&D Center, http://trace.wisc.edu/world/doc_access/

Turvey, M.T., and Carello, C., 1995, Dynamic touch, In W. Epstein and S. Rogers (Eds.), Perception of space and motion, pp. 401–490, Academic Press, San Diego, CA

Ungar, S., Simpson, A., and Blades, M., 2004, Strategies for organizing information while learning a map by blind and sighted people, In S. Ballesteros Jiménez and M.A. Heller (Eds.), Touch, Blindness and Neuroscience, pp. 271–280, UNED Press, Madrid, Spain

van Erp, J.B.F., Carter, J., and Andrew, I., 2006, ISO's work on tactile and haptic interaction guidelines, Eurohaptics Int. Conf., EH 2006, July 3–6, 2006, Paris, France, Procs., pp. 467–470

Vega-Bermudez, F., Johnson, K., and Hsiao, S.S., 1991, Human tactile pattern recognition: Active versus passive touch, velocity effects, and pattern of confusion, J. Neurophysiology, Vol. 65, pp. 531–546

VIEWPLUS, undated, http://www.viewplus.com/products/touch-audio-learning/

VTPL, undated, http://www.virtouch2.com

Wall, S.A., and Brewster, S., 2004, Providing external memory aids in haptic visualizations for blind computer users, Procs. Fifth International Conf. on Disability, Virtual Reality and Associated Technology, 20th–22nd Sept. 2004, Oxford, United Kingdom, pp. 157–164, University of Reading, Reading, UK, Also available at http://www.icdvrat.reading.ac.uk/2004/index.htm

Wall, S.A., and Brewster, S., 2006, Feeling what you hear: Tactile feedback for navigation of audio graphs, In Procs. ACM CHI 2006 (Montreal, Canada), ACM Press Addison-Wesley, pp. 1123–1132

Wall, S.A., Paynter, K., Shillito, A.M., Wright, M., and Scali, S., 2002, The effect of haptic feedback and stereo graphics in a 3D target acquisition task, In S.A. Wall, B. Riedel, A. Crossan and M.R. McGee (Eds.), Eurohaptics 2002, Conf. Procs, pp. 23–29, University of Edinburgh, Edinburgh, UK

White, B.W., Saunders, F.A., Scadden, L., Bach-y-Rita, P., and Collins, C.C., 1970, Seeing with the skin, Perception and Psychophysics, Vol. 7, pp. 23–27

Whitstock, R.H., Franck, L., and Haneline, R., 1997, Dog guides, In B.B. Blash, W.R. Wiener, and R.L. Welsh (Eds.), Foundations of orientation and mobility, 2nd. Edn., pp. 260–283, AFB Press, New York

Wies, E.F., Gardner, J.A., O'Modhrain, S., and Bulatov, V.L., 2001, Web-based touch display for accessible science education, In S. A. Brewster and R. Murray-Smith (Eds.), Haptic human-computer interaction, pp. 52–60, Springer, Berlin, Germany

Wing, A.M., Haggard, P., and Flanagan, J.R. (Eds.), 1996, Hand and brain: The neurophysiology and psychology of hand movements, Academic Press, San Diego, CA, USA

Yu, W., Ramloll, R., and Brewster, S., 2001, Haptic graphs for blind computer users, In S. Brewster and R. Murray-Smith (Eds.), Haptic human-computer interaction, pp. 41–51, Springer, Berlin

Zimmerman, G.J., and Roman, C.A., 1997, Services for children and adults: Standard program design, In B.B. Blash, W.R. Wiener and R.L. Welsh (Eds.), Foundations of orientation and mobility, 2nd. Edn., pp. 383–406, AFB Press, New York

5 Mobility: An Overview

Learning Objectives

Most existing urban environments were not designed using modern accessibility principles and consequently many visually impaired and blind people experience severe difficulties in travelling even short distances in these public spaces. Although a number of blind and visually impaired people are confident and experienced travellers, significantly greater numbers of blind and deafblind people feel unable to go out without a guide. Thus, the development of travel aids, which could be used to support independent travel by blind, visually impaired and deafblind people, is an important application area for assistive technology.

Although there has been a reasonable amount of work in this area and a number of systems have been developed, many have not gone beyond the prototype stage. Of the others, some have totally disappeared and others are used by relatively small numbers of blind and visually impaired people. Therefore, there is a need for the development of reliable travel aids which have good technical performance and are also attractive to the end-user community. However, this will only be a partial solution, since the widespread adoption of fully accessible environments combined with the development of user friendly integrated assistive technology for mobility will be required to remove the barriers to mobility currently experienced by blind and visually impaired people. This chapter provides an overview of the current state of the art as an introduction to the more detailed discussion of a number of these devices and developments in subsequent chapters.

The learning objectives for this chapter include the following:

- Understanding the different types of barriers to independent mobility experienced by visually impaired and blind people.
- Understanding the approaches used by blind and visually impaired people to support (independent) mobility.
- Appreciating the range of assistive technology systems for mobility that have been developed and the general principles of several of these systems.
- Understanding the concept of a fully accessible environment and an overview of the progress made towards implementation.

5.1 Introduction

Much of everyday life involves the activity of moving from one place to another. Unfortunately, in many countries, facilities such as shops, leisure and sports centres are increasingly being located outside the centres of population. For visually impaired and blind people, most of whom are not able to drive, access to such locations depends on the availability of public transport and, in many cases, the assistance of a human guide. In addition, there is a tendency for people to live away from their workplace and, at least in the industrialised countries, increasing numbers of people are travelling abroad on holiday. Thus, the ability to travel at least short and medium distances and find one's way around public spaces and commercial centres is important for personal independence, for employment and for participating in shopping and leisure activities. Consequently, visually impaired and blind people need assistive technology systems to support effective travel and to contribute to independent living and working. Further, the design of urban and other environments should be improved to make them easy to move around in and obstacles that could endanger safety, particularly of visually impaired and blind people, should be removed. This is likely to have benefits for everyone, whether visually impaired, blind or sighted.

Travelling, even for relatively simply trips, involves a number of different activities which are most easily carried out using vision. Consider a journey across a city by bus to visit a hospital consultant. The tasks involved include the following:

- Being able to avoid obstacles on the pavement.
- Walking in the right direction.
- Crossing the road safely.
- Finding the correct bus stop.
- Knowing which is the right bus.
- Paying the correct fare.
- Finding a vacant seat.
- Knowing when to get off the bus.
- Crossing the road safely (at a different location and probably using a different type of crossing).
- Walking to the hospital entrance.
- Finding the main reception desk.
- Finding and using a lift (elevator) to the correct floor.
- Locating the waiting room and the consultant's room.
- Leaving the hospital and repeating the travel process in reverse to return home.

For such tasks, sighted and hearing people mainly rely on their sense of sight, complemented by hearing. Visually impaired and blind people are only able to use their sense of sight to a limited extent or possibly not at all and, except in the case of deafblind people, will be largely reliant on hearing. Both blind and deafblind people therefore require support from assistive technology to carry out many of these travel activity tasks. As can also be seen from the above typical journey, there is a need for the design and layout of accessible urban environments, accessible public spaces and buildings and accessible signage to make independent mobility much easier for blind and visually impaired people. This chapter presents an overview of assistive technology systems for these two important topics: travel and environmental accessibility for visually impaired and blind people.

5.2 The Travel Activity

Even simple travel activities, such as a bus journey across a city to visit a hospital or sports centre or the daily commute by suburban railway or metro, generate a fairly long list of travel subtasks (Harper and Green 2002). A simple classification of these tasks results in two main categories—mobility and environmental access. These categories and their further subdivisions are illustrated in Figure 5.1.

Independent mobility is a multifaceted topic and has been researched from a number of different disciplinary perspectives, including that of psychology. Combining the different basic research perspectives with that of the end-user can contribute both to an understanding of the subject and to the development of successful assistive technology systems to support independent mobility. However, leaving out the end-user perspective, as has frequently happened, is likely to lead to systems which are not taken up by end-users. This basic research can then be used in assistive technology systems for the mobility component of travel. Embedding assistive technology systems into the urban environment is a more recent development but is becoming more commonplace. Such environmental assistive technology systems facilitate access to information and can increase safety.

5.2.1 Understanding Mobility

Investigation of the spatial understanding of blind people dates as far back as the late seventeenth century, for example, the correspondence between the scientist

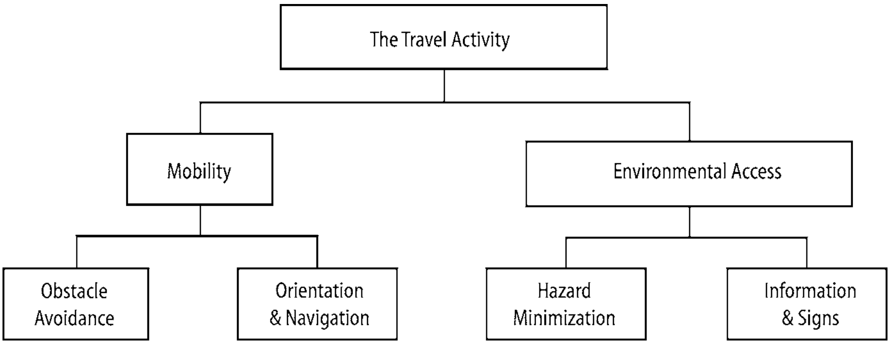


Figure 5.1. Tasks within the travel activity

Molyneux and the philosopher Locke on this subject. However, it is only relatively recently that empirical studies have been performed and their results used in the design of assistive technology systems for mobility. Research in this area is still continuing. In this section, a study by Ungar (2004) and a model by Brambring (1985) are reviewed to give a flavour of the insight into mobility that can be obtained from empirical and theoretical research. Before discussing these studies, some concepts and definitions used in mobility research will be presented:

1. *Near-space and far-space.* Near-space is the space immediately around the person's body (or their body plus a short-range assistive device such as a long cane). This space can be explored by touch and thus it is often called the haptic-space. Far-space is distant geographical space, information about which is required for travel; hence it is also sometimes referred to as the locomotor space. Travelling around a campus or across a city are examples of a far-space or locomotor-space problem. Understanding how mobility mechanisms work in far-space is a key ingredient for travel success.
2. *Past experience and new experience.* Using past experience involves storing spatial information about a journey and recalling this information when the journey is being repeated. Preparing for new experiences involves applying knowledge from prior experience to new journeys or spatial tasks. In empirical studies it is often important to make a distinction between these two types of situations, one based on memory and the other based on inference.
3. *History of visual impairment.* Empirical studies, particularly in psychology, frequently involve a variable relating to the onset of visual impairment or blindness. Congenitally or early blind or visually impaired people may have different cognitive approaches and travel strategies from people who became blind or visually impaired later in life. Late blind people have the advantage of familiarity with visual concepts and visual knowledge of a (large) number of routes, but may experience problems developing new travel strategies. Early or congenitally blind people have the advantage of being accustomed to relying on other sources of information than sight. Some of them also have well developed strategies for independent travel as a blind or visually impaired person. Therefore, the existence and extent of prior visual experience can have a significant effect on the performance of mobility tasks in empirical studies. Even more significantly, it can also affect end-user requirements and satisfaction with assistive technology.
4. *Body-centred and external referencing strategies.* There are two main approaches to fixing the location of an object in (haptic) space. Its location can be referenced either to the body (and/or movements) of the person searching the space or to other objects in the haptic space. The two approaches are called body-centred and/or movement centred and external referencing. The term ego-centred referencing is sometimes used for the first approach.
5. *Cognitive maps.* A spatial pattern of important locations or way-finding points is often described using a physical map or diagram. A cognitive map is a mental version of this coding or mapping activity. It involves mentally storing the

ordering or relationships between a number of variables that can be physical objects such as way-finding points or abstract objects such as ideas.

The findings of empirical investigations (Ungar 2004) of mobility and its operating mechanisms for visually impaired and blind people included the following two outcomes:

1. *Mobility in near-space*. In so-called "table-top" experiments where the participants demonstrate mapping abilities for a tableau of objects within arm's reach, the results indicated that body-centred mapping strategies are used by blind people who are congenitally or early blind. Deeper investigative studies have shown that these strategies are less robust in more demanding tasks that require extrapolation; in these more complex tasks the use of external references was found to be more successful.
2. *Mobility in far-space*. These studies were performed in a larger environment, typically, school grounds, offices or constructed layouts, and frequently with children of school age. One clear outcome was "that the lack of visual experience does not prevent the acquisition of spatial representation" (Ungar, 2004). However, the studies seemed to show that blind people tend to use body-centred (egocentric) rather than external referencing spatial strategies to record their experiences of far-space and that external referencing strategies can be more robust and successful for the more complicated mobility tasks in far-spaces. Ungar (2004) has also considered the activities that could encourage an improved appreciation of the layout of far-space by visually impaired and blind people. He suggests more specialised mobility training in the use of all the senses from an early age and the use of tactile maps.

Travelling in the real world is considerably more complex than mobility in a limited test environment. A model that is frequently cited is that due to Brambring (1985), as shown in Figure 5.2.

In Brambring's travel model, the locomotion of a blind person is considered to comprise two types of processes, one concerned with perception and the other with orientation. Perception is interpreted in the most general sense. In this context, it refers to a blind person obtaining information about the environment through any of their senses, whilst orientation is the knowledge of general position on the journey path. With this classification, the two areas of perception and two areas of orientation are refined as follows:

- *Perception and obstacle detection*. The detection of obstacles means the perception of potentially hazardous objects in the environment ahead of time so they can be avoided. This has always been considered a primary requirement for aided mobility. This is a perception activity in the *near* portion of locomotor space.
- *Perception and the identification of landmarks*. Blind travellers have been found to use a greater variety of and different types of landmarks than sighted people in order to determine their location with respect to specific points on a route. The class of landmarks used by blind travellers includes subtle environmental

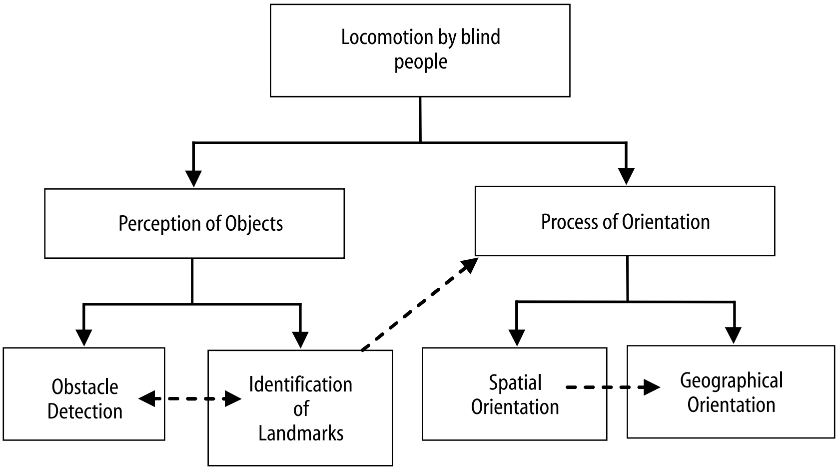


Figure 5.2. Brambring's decomposition of travel by blind people

clues, such as rises and falls in the path and changes in the texture of the pathway, the presence of walls and hedges, changes in the type and pattern of nearby objects like railings, sound landmarks, including traffic sounds and traffic crossing alerts, as well as odours and temperature changes. Other than intense odours and loud sounds, which can be perceived at a distance, most of the perceptual activity of a profoundly blind person is carried out in the *near* portion of locomotor space. In addition, a large number of landmarks are required to learn a route. With increasing quality of vision, the class of landmarks approaches the class of landmarks used by sighted people, namely street signs, significant buildings (particularly pubs!) and junctions. In this case, the perception activity moves into the *far* space. But it should also be noted that there are considerable differences in the type of landmarks used by different sighted people. Differences in the types of landmarks used by different blind people are likely to be a consequence of differences in individual perceptive abilities, rather than differences in cognitive processes or other factors, as in the case of sighted people. This is due to the need for profoundly blind people, in particular, to register as many landmarks as possible and therefore to take note of all the landmarks they can perceive. In some cases, blind (and sighted) people may use several landmarks at a time to confirm location.

- *Spatial orientation*. This is a near-space activity, and is the ability to maintain a position in the immediate environment of the traveller. Spatial orientation issues include the optimal position on a pavement and the ability to walk without veering from a desired course.
- *Geographic orientation*. In contrast to spatial orientation, geographic orientation is a far-space activity. It is the ability to determine position in the geographical space of the entire journey. Geographical orientation is dominated by informa-

tion and strategy about how to get from place A to place B. *Navigation* is the term often used in this context.

5.2.2 Assistive Technology Systems for the Travel Process

There have been considerable advances in the available approaches to developing assistive technology systems. In particular, there is increasing recognition of the importance of the involvement of end-users in all stages of the design process and increasing use of human-centred, user-centred and participative design approaches. Assistive technology researchers can also draw on the social model of disability and the framework for modelling assistive technology presented in Chapter 1. In the past there has been a tendency to focus on technical aspects of travel by blind and visually impaired people and ignore the human dimension. This led to the application of (new) advanced technologies without consideration of the wider context. Success in developing (electronic) travel aids for blind and visually impaired people has been rather limited. Many of the devices have not gone beyond the prototype stage. Only relatively small numbers of the most successful electronic devices are in use and the long cane and guide dog (see Sections 5.4.1 and 5.5.1) are still the most widely used travel aids for blind and visually impaired people. This indicates a need for more research into the reasons for this lack of success to target future research and development in this area better. Possible reasons (which would require experimental verification) include the following:

- Excessive complexity, making devices difficult to learn and use, combined with a lack of training.
- High costs and a lack of research and development funding.
- Inappropriate appearance.
- Not meeting users' needs.
- Awkward or heavy to carry.
- Not providing significantly greater functionality than the long cane.

From the user's point of view, it is clearly advantageous to have one device that supports all travel activities. From the engineering design perspective, it is useful to have a categorisation of the different types of activities involved. This can be motivated by considering typical travel activities, for instance, travelling down an urban street (Figure 5.3).

Typical travel issues that can be identified from the street scene are as follows.

Obstacle avoidance

- Clear-path ahead: the need to have a clear forward path at chest height, at leg height (litter bin) at head height (tree branches).
- The need to have sufficient space at the sides to proceed: for instance, scaffolding or a wall at one side, a doorway or arch or a path that narrows need to be identified.

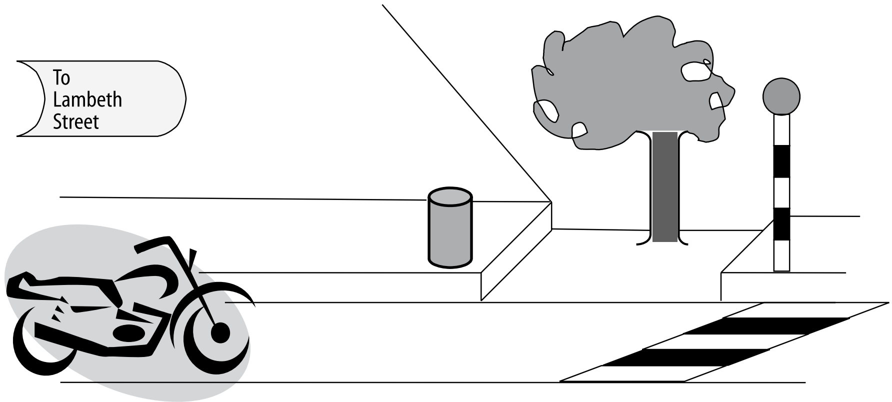


Figure 5.3. Obstacles, hazards and information in an urban street scene

- Path level changes: the need to identify path drop-offs, sudden path step-ups and down slopes and inclines.

Mobility: orientation and navigation

- Safe pavement position, safe path following.
- Safe traffic crossing for pedestrians.
- Route finding: finding and following the desired route.
- Indication that the destination has been reached or a desired object, such as a door or waste bin, found.

Environmental access: hazard minimization

- Safe traffic crossing for pedestrians.
- Indicators, for instance, of the locations of the ends of pavements and traffic crossings.
- Street furniture being sited to avoid it becoming an obstacle and hence a hazard to pedestrians.

Environmental access: information and signs

- Information about location, access to street signage, public transport information and other public notices.

Travel for blind people can be facilitated by both the use of assistive technology and accessible environmental design. For instance, street furniture, such as bins and lampposts, should be positioned so that it does not present an obstacle. However, assistive devices will still be required, to detect people and any street furniture that still presents an obstruction. In general, environmental access will be promoted by accessible design, whereas both accessible design and assistive technology will support mobility.

Although some attention is now being given to an integrated approach to the development of assistive technology and the overcoming of infrastructural barriers, this has not yet had a significant impact on work in the area. Therefore, most researchers are still focussing on the development of assistive devices for either obstacle avoidance or orientation whilst there is a separate body of work on removing barriers and making environments more accessible to blind and visually impaired people.

Most mobility (obstacle avoidance) and orientation devices carry out the following two functions:

1. Obtaining near-space or far-space information
2. Providing this information to the user in an appropriate form

Most of the research in the area has concentrated on different methods for obtaining information and there has been less research effort on communicating this information to the end-user. This may be one of the reasons why a number of devices have been unsuccessful. Communicating information to the end-user raises the following issues, many of which still require further investigation:

- The choice of sensory modality, that is, touch or hearing. For some visually impaired people it might be possible to process the information to enable them to receive it visually. For instance, this could include presenting information so that it could be detected by central vision and therefore used by visually impaired people who only have central vision.
- The provision of information in a way that does not interfere with the use of sensory environmental information. For instance, speech or sounds should be provided by an ear phone to one ear only, so as not to impede the detection of auditory environmental cues.
- The need for training, including in the processing of complex tactile or auditory information and whether this training is required from an early age. Currently the processing capabilities of touch and hearing limit the information that can be provided. However, it is possible that appropriate training, particularly if carried out at an early age, would overcome this limitation or at least reduce its impact.
- The amount of information provided. Most existing devices provide basic information, for instance on whether or not there is an obstacle, rather than complex information which would give the user an overview of the scene.

A consequence of this type of simple decomposition of the travel task is that in the past, assistive technology engineers have devised *ad hoc* solutions to some of these individual problems in the separate categories: (1) obstacle avoidance assistive technology and (2) orientation and navigation assistive technology. It is only in more recent years that the idea of taking an integrated approach to the development of assistive technology solutions has been proposed and in some cases implemented.

5.2.2.1 Categorisation of Travel Aids

There are a number of different categorisations of travel aids, which can be stated as follows:

1. Into primary and secondary aids based on whether the device can be used on its own or is used to supplement another device:
   - *Primary aids*. These are used on their own to deliver safe mobility. A typical example is the long cane.
   - *Secondary aids*. These devices are used to supplement a primary device. However, they have not been designed to deliver safe mobility when used on their own. A hand-held ultrasonic torch device for obstacle detection is an example of a secondary aid that could be used to augment the use of a long cane.
2. Based on the functionality of the device:
   - Mobility devices that support obstacle avoidance.
   - Orientation and navigation devices that provide information on landmarks and support route finding.
   - Environmental access assistive technology.
   - Devices which support object finding.
3. Based on the technology used to obtain the environmental information:
   - Ultrasonic.
   - Infrared.
   - Camera.
   - Global positioning system (GPS).
   - Mobile phone technology.
4. Based on the way information is provided to the user:
   - Tactile, generally by vibration.
   - Speech.
   - Sounds of varying loudness and pitch.
   - Musical tones.
5. Based on how the device is carried:
   - Cane.
   - Other hand-held device.
   - Carried in pocket.
   - Carried in back pack.

To give a uniform presentation to the many different assistive technology devices and applications it is useful to use the comprehensive assistive technology (CAT) model. This provides a very general framework for the development and comprehension of assistive technology. The assistive technology system block diagram is

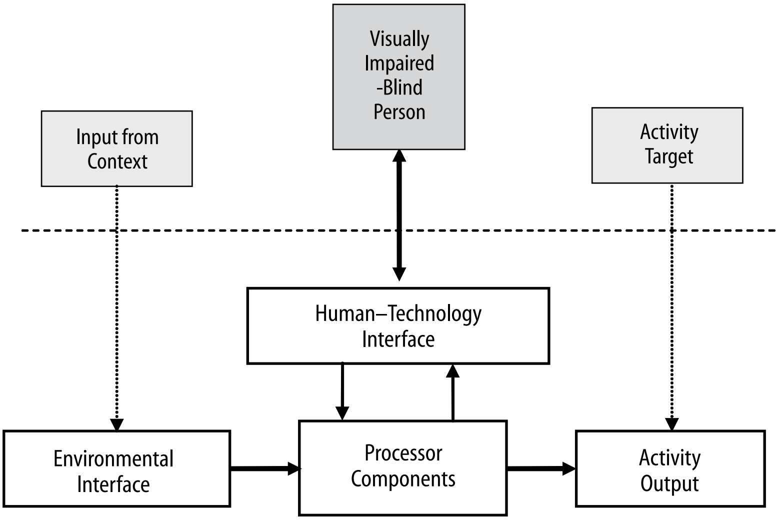


Figure 5.4. Assistive technology system block diagram

shown in Figure 5.4 and this is usually specialised for the particular system under discussion. Further background material on the CAT model is given in Chapter 1 where it is introduced and explained.

5.3 The Historical Development of Travel Aids for Visually Impaired and Blind People

Historically visually impaired people have used a number of different methods and tools to support mobility. From the earliest times blind and visually impaired people have used sighted guides, trained animals and simple devices, such as sticks, to give them varying levels of mobility. However, the development of assistive technology systems that will enable all blind people to travel with the same confidence as sighted people is a complex multi-factor problem.

For centuries, visually impaired people have used some form of stick for probing when travelling. However, the long cane, which is currently the most popular device, did not come into widespread usage until around 1950. Guide dogs were first used in the 1920s in Austria (see Figure 5.5).

The first electronic mobility aid, Noiszewski's Elektrftalm, was developed in 1897. However, serious developments occurred after the Second World War and through the 1950s and 1960s. Once the potential applications of remote sensing by using ultrasound and infrared radiation were realised, research effort was directed at obtaining environmental information using these technologies. Early ultrasonic mobility devices include the Franklin Institute Electronic Cane, Mowat Sensor, and the Pathsounder. Throughout the 1960s and 1970s, obstacle detection devices continued to be developed using a variety of sensing technologies. The development of lasers allowed the so-called "Laser canes" to be produced (see Figure 5.5). Table 5.1 shows a survey of the details for a number of these obstacle detection devices.

In the last decade there has been increased interest in devices to support orientation and spatial sensing. The systems developed include 'Talking Signs' and 'Sound Buoys', which transmit a remote signal, for instance using infrared radiation. When a traveller moves into the vicinity of the device, the device gives the user an audible warning. These systems have proved to be successful, but installation can be expensive. There are currently many different research projects worldwide looking into different forms of access to mobility and orientation information for visually impaired people. These involve technologies and applications as diverse as ultrasonic systems, infrared systems, GPS technology, sound signs, vision substitution systems, vision enhancement technologies, cell telephony and computer network technologies.

To date assistive technology systems for mobility, orientation and navigation have been developed largely independently of each other. Consequently, users often require separate devices to carry out different parts of the travel process. In addition, some of the devices are quite cumbersome, which has probably contributed to poor take-up. However, advances in signal processing, information and communications technologies, ultrasonic technology and device miniaturisation make the development of a small size multi-purpose assistive technology system a realistic possibility. However, this will require a change in the direction of research effort to focus on the development of a holistic travel system for visually impaired and blind people. This system would have both obstacle avoidance and orientation capability and would incorporate interfaces to the modified built environment to remove barriers to mobility.

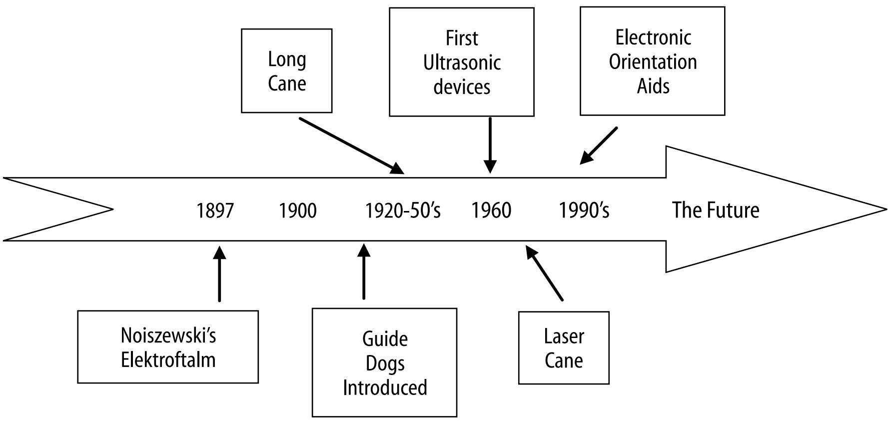


Figure 5.5. Historical trends for mobility devices (Li 2001)

Table 5.1. Survey of obstacle detectors

Device name	Approx. year	Technology	Body position	References and notes	Current availability
Sonicguide	*ca.* 1965	Ultrasonic	Head worn	Kay 1974, 1980; Jacobson, 1979	Discontinued
Bat K sonar cane	*ca.* 1965	Ultrasonic	Clips onto long cane	BFB 2006	Commercially available
Pathsounder	*ca.* 1966	Ultrasonic	Neck worn	Russell , 1966	Not available
Mims device	*ca.* 1970	Infrared light	Head worn	Mims 1972	Prototype only
Mowat sensor	*ca.* 1970	Ultrasonic	Hand-held, torch-like	Farmer and Smith, 1997	Not known
Laser Cane™	*ca.* 1970	Infrared system	Hand-held cane	Benjamin, 1968; Benjamin *et al.* 1973; Blasch *et al.* 1989a,b; NR 2006	Commercially available
Polaron	*ca.* 1980	Ultrasonic and ranging (sonar) system	Hand-held torch format or chest position worn	NR 2006	Commercially available
Sensory 6	*ca.* 1980	Ultrasonic and ranging (sonar) system	Head worn	Farmer and Smith 1997; BRY 2006	Custom orders
Wheelchair Pathfinder	*ca.* 1980	Ultrasonic and Infrared systems	Wheelchair mounted	NR 2006	Commercially available
Sonic Pathfinder	*ca.* 1990	Ultrasonic	Head worn	Farmer and Smith 1997; Heyes 1993	Not known
Ultra Cane™	*ca.* 1990	Ultrasonic	Hand-held cane	See Chapter 6 of this volume and the references therein; UC 2006	Commercially available
Tom Pouce	*ca.* 1990	Infrared system	Cane mounted	Farcy *et al.* 2006	Commercially available
Teletact	*ca.* 1990	Infrared system	Hand-held/Cane mounted	Farcy *et al.* 2006	Commercially available
Miniguide	*ca.* 1998	Ultrasonic	Hand-held	GD 2006; GDP 2006	Commercially available

5.4 Obstacle Avoidance AT: Guide Dogs and Robotic Guide Walkers

5.4.1 Guide Dogs

Guide dogs have been used as mobility aids for blind and partially-sighted people for over 75 years in the UK and elsewhere. The usual breeds are Labradors, retrievers (and crosses between those breeds) and German shepherds, although a small number of other breeds are also used. Providing freedom and independence, a guide dog can support safe travel for a visually impaired person and give them the confidence to go out on their own with the dog, without needing assistance from a sighted guide. The dog's innate intelligence means that its performance in many circumstances can approach that of a sighted guide. Figure 5.6a shows a guide dog in action in an urban environment. Note also the long cane being used by another pedestrian in the figure.

Both the guide dog and owner are trained by experienced professionals. For example, in the UK, the charity The Guide Dogs for the Blind Association (Guide Dogs) provides such training. A properly trained dog can guide its human companion with as much safety as a sighted guide. The dog can be trained to stop for dangerous obstructions and approach up and down curbs with timed precision. The dog will slow down while walking over rough, broken or difficult surface areas. Guide dogs are trained to learn to recognise frequently travelled routes, using their intelligence in everyday situations such as crossing the road. A guide dog will stop if the way ahead is blocked with dangerous features or obstacles. Many guide dogs can also obey commands, for instance, to find the train doorway on the metro or underground (see Figure 5.6b).

Factors such as whether or not a visually impaired person is comfortable with dogs, and whether they are prepared to devote sufficient time to grooming, feeding, exercising and cleaning up after one will determine whether the guide dog is the right mobility choice for them. In the UK, it costs £10 a day to breed, train and support each of the 4700 guide dogs currently in service and this cost, including food and vet bills, is covered by the Guide Dogs charity. Members of the public are discouraged from distracting a working guide dog by petting or feeding it; however, if the owner requires help, they will indicate this by lowering the handle on the harness. The working life of a guide dog is around six-and-a-half years and once the guide dog's service years are over there are many people who wish to provide a happy retirement home for a guide dog. The person will then have to be retrained with a new dog.

It is useful to analyse how a guide dog is effective using the assistive technology system block diagram as shown in Figure 5.7. This gives an interesting insight into the features that an engineering assistive technology *guide* would have to have to be a success.

a

b

Figure 5.6a,b. Guide dogs in action: **a** guide dog and long cane mobility aids in action in a pedestrian area; **b** guide dog leading a person onto a train (photographs reproduced by kind permission of the Guide Dogs for the Blind Association, UK)

5.4.2 Robotic Guides and Walkers

The combined visually impaired person and guide dog system is an example of a shared control system. The person and the dog interact with one another to achieve the activity of guided and safe mobility. Designing and implementing an *engineered* shared control assistive technology system that replicates some of the functions of a guide dog is a challenging problem, but prototypes do exist as exemplified by the GuideCane (Borenstein and Ulrich 1997; Ulrich and Borenstein 2001) and the Robotic Cane (Aigner and McCarragher 1999).

The ideas underlying a robotic walker include the following:

- Real-time analysis of the environment and then computing and following an appropriate optimal travel direction that avoids any obstacles in the user's path.

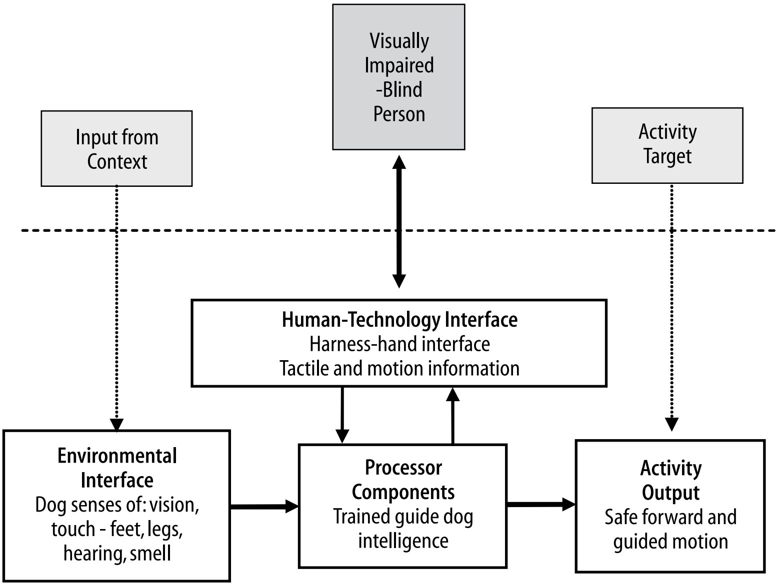


Figure 5.7. Assistive technology system block diagram – guide dog

- The user intuitively reacts to changes in the direction of the cane handle caused by the movement of the wheelbase and is thereby steered along the desired path.
- The physical interface between the user and the device is the handle. The angle of the handle is adjustable for users of different heights. The handle has a pointing device that is used to indicate the desired direction of travel relative to the cane's current orientation.
- The device should be light and compact to make it easy to use, including lifting into public transport and using on stairs. The electronic components should require minimal power to minimise battery weight.
- The system should be as natural and intuitive to use as possible, and should require minimal training input.
- The device should either be unobtrusive or have an attractive appearance. For younger users a science fiction/robot look might work, but this might not be appropriate for older users.

The GuideCane is shown in Figure 5.8 and a description of some of its features follow.

As with a long cane, the user holds the GuideCane in front and follows the handle. The device is considerably heavier than a long cane, but it rolls on wheels that support the weight. The wheels can be steered to the left or right relative to the cane by a servomotor controlled by the built-in computer. Encoders on the wheels determine their relative direction. The user can indicate a desired direction of

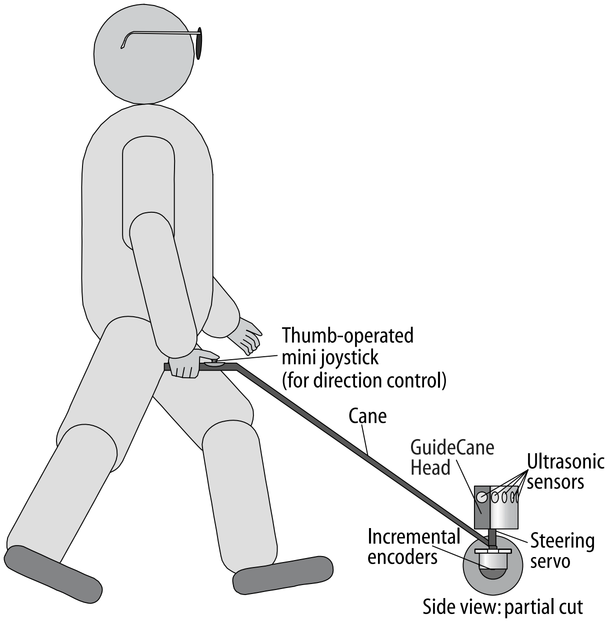


Figure 5.8. GuideCane and pedestrian user (reproduced by kind permission of Professor J. Borenstein, University of Michigan, USA)

motion by pressing a mini joystick in the handle in that direction. A new direction is usually indicated before the change of direction actually occurs.

There are ten ultrasonic sensors for obstacle detection. They can detect any obstacle in a 120° wide sector ahead of the user. The built-in computer uses this sensor data together with the user input and data from the encoders to determine the direction of travel. If there is an obstacle in the desired travel direction, the obstacle avoidance algorithm provides an alternative path around the obstacle and then resumes in the desired direction.

When the wheels turn sideways to avoid an obstacle, the user almost automatically changes orientation to follow the cane, so the walking trajectory follows that of the cane. Once past the obstacle, the user can continue to walk along the new line of travel or use the GuideCane's dead reckoning capability to return to the original line of travel. Figure 5.9 shows these steps.

When a step down is reached, the wheels drop off the edge until the shock-absorbing wheelbase hits the step. Since the user walks behind the cane and this signal is very strong, the user has sufficient time to stop. A step up could be detected by incorporating additional front facing sonars, but this was not implemented in the first prototype.

The Guidecane has not gone beyond the prototype stage. However, one of the aspects of the existing design, which could act as a barrier to its widespread use if not modified, is its appearance and size. A block diagram of the Guidecane system is given in Figure 5.10, where it can be seen that the interfaces required are quite

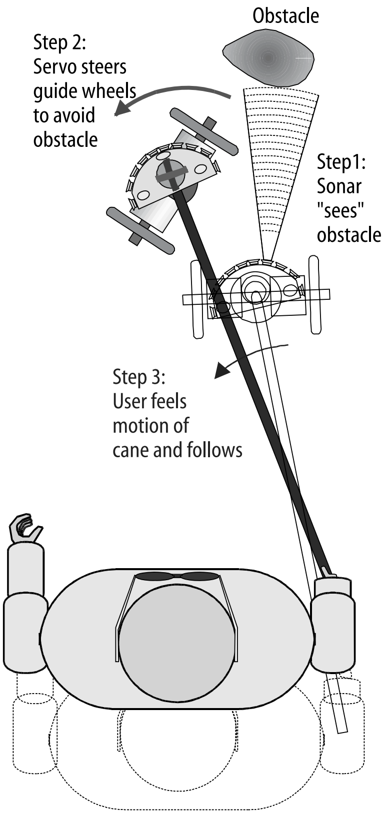


Figure 5.9. GuideCane – obstacle avoidance manoeuvre (reproduced by kind permission of Professor J. Borenstein, University of Michigan, USA)

complex and the processor includes a small computer running obstacle detection algorithms and route planning routines.

A key design problem with the robotic walker concept is deciding how control between the user and the walker is shared. It is important that the user is able to override the walker but not in conditions that might cause injury to the user if they continued in motion. This is exactly the control sharing that is observed between a user and their guide dog. If the guide dog does not like the forward path, this information is indicated to the user by the dog's hesitancy or by the dog simply coming to a halt. However, there is the important difference between a guide dog and robotic walker system that a guide dog possesses intelligence and that some degree of 'discussion' with the system user, its human companion, is possible.

In the case of the GuideCane, the robotic walker is a semi-autonomous system, with full autonomy for local navigation (obstacle avoidance) but it uses the travelling pedestrian's skills for global navigation (path planning and localisation). The main task for the robotic walker is to steer round obstacles and proceed in the desired travel direction. Thus, its performance is directly related to the per-

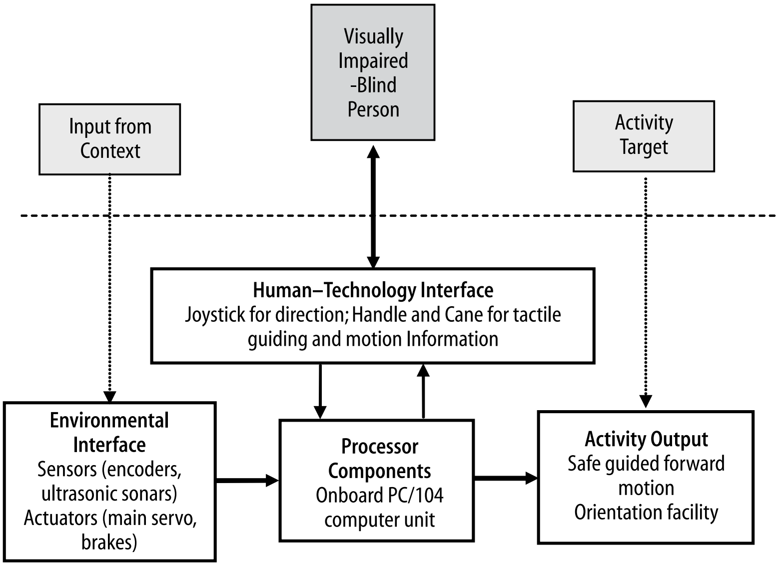


Figure 5.10. Assistive technology system block diagram – GuideCane

formance of the obstacle avoidance algorithm and to achieve safe travelling at fast walking speeds through cluttered and unknown environments, mobile robot obstacle avoidance algorithms and methods are used.

5.5 Obstacle Avoidance AT: Canes

There are three different types of canes:

1. Symbol canes
2. Long canes
3. Technology canes

Symbol canes

The symbol cane is a lightweight cane of length 1.0–1.5 m. It is slim, cylindrical, white and often made of aluminium. This cane is simply an indicator of the fact that the cane-bearer has poor sight. It is a useful signifier to other pedestrians, drivers, and people in public and commercial settings that its bearer is visually impaired. The symbol cane is *not* designed to be used as a tactile detector of obstacles.

A more robust version of the symbol cane is available and often called a guide cane or mobility cane. Guides canes are also indicator canes but have a more rugged construction to allow them to survive urban wear and tear longer than the usual slim symbol cane.

5.5.1 Long Canes

The long cane is the standard mechanical mobility device used by many visually impaired and blind people. It should be selected to stand to the height of the breastbone of the user. The cane is of robust construction with a purposely-designed handle-grip and tip. The tip of the cane can be either a roller ball or just a plain blunt tip. With a roller ball, the ball remains in contact with the ground and the cane is “bounced” from side to side, to sense any obstacles. The shank of the cane transmits tactile obstacle detection information to the user. If the cane has a plain tip then the cane is “bounced” from side to side so as to make contact with the ground at the two outermost points of its trajectory. The resulting tactile information on obstacle locations is transmitted by the shank of the cane to the user. The long cane has for a long time been considered to be the most effective and efficient mobility aid devised for safe and independent travel for blind people. It will be interesting to see whether the more recently developed electronic travel aids, such as the ultracane and the teletact, displace it from this ranking.

The long cane is a simple, robust, reliable and inexpensive device that provides immediate tactile information about the path of travel in front of the user by scanning. It also gives some information about the surface and the condition of the ground ahead, as well as some sound information (echolocation cues) about the path ahead. Users can relate the sound and tactile information to their memories of previous trips over the same route. This provides limited but possibly important information about navigation or orientation on longer journeys. Due to its robust construction, the long cane can be used in poor weather. It requires no maintenance and the replacement accessories are small, inexpensive and only required infrequently. The long cane will give warnings of drop-offs and objects ahead and thereby provide protection for the lower part of the body from collisions. However, it is unable to provide information about overhead obstacles to prevent collisions to the upper part of the body.

The long cane is instantly recognised by sighted pedestrians, who tend to move out of the way and give the user a clear-path. However, this means that the long cane overtly labels the user as blind and some visually impaired and blind people might prefer to avoid this labelling. Other disadvantages include that the long cane can be quite tiring to use over a long distance. Arm fatigue is common and the double touch cane technique is more tiring than the roller ball method. The long cane often snags holes and cracks in poorly finished pavements and this usually slows the pedestrian’s walking pace considerably. One example of obstacle snagging occurs when negotiating an open pedestrian area, such as a public square, where the long cane can easily slide under a bench that is across the direction of travel. Learning to use the long cane requires special training because incorrect use of the cane can be dangerous to both the user and possibly to others. The length of the training period depends on the mode of use of the cane and the individual’s capabilities and perseverance. Figure 5.11 shows a block diagram for the long cane system.

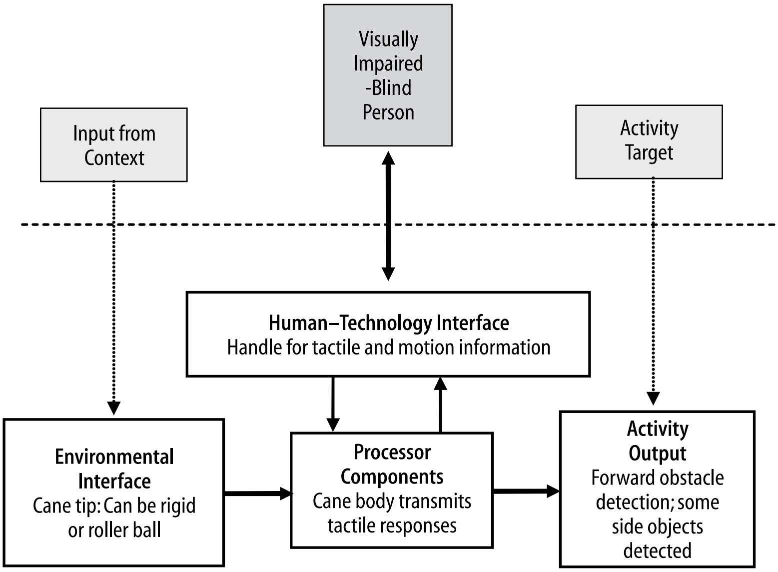


Figure 5.11. Assistive technology system block diagram – long cane technology

5.5.2 Technology Canes

The term technology cane is used for a group of obstacle detection canes that are constructed on long cane principles but use additional technology to detect obstacles and relay information about the obstacles to the cane bearer. Figure 5.12 shows the basic construction of a technology cane.

The full complexity of a modern technology cane is captured in considerable detail using the block diagram shown in Figure 5.13.

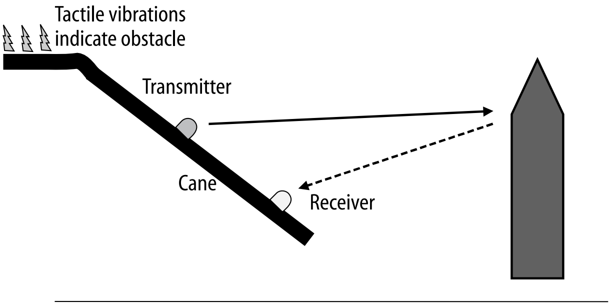


Figure 5.12. Generic technology cane construction

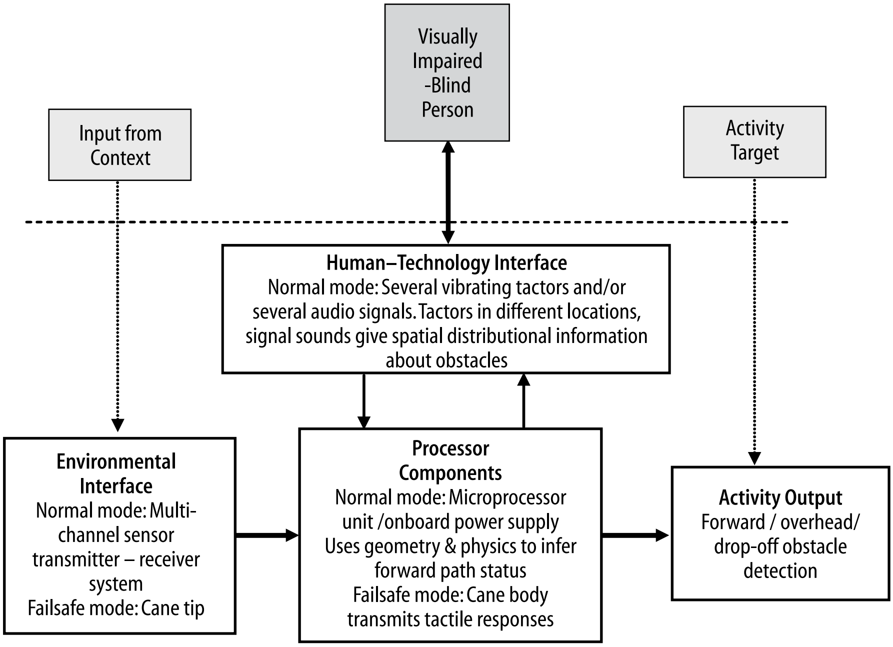


Figure 5.13. Assistive technology system block diagram – obstacle detection technology

The following two main technologies are used in the obstacle detection component of these modern technology canes:

- Infrared emission and detection systems, which are generally referred to as "laser canes" (Light Amplification by Stimulated Emission of Radiation), use coherent beams of infrared radiation concentrated on a single frequency.
- Ultrasound emission and detection systems. In Chapter 6, a full presentation of the development of the UltraCane™, which uses ultrasound, is given by the developers of this particular advanced technology cane.

A brief review of both "laser" cane and ultrasound systems is presented below.

Laser cane systems

The use of light in an obstacle detection system dates back to 1943 when incandescent light was first used in prototype systems. Today, infrared lasers are generally used. For example, a laser cane might use miniature solid-state gallium-arsenide room temperature injection lasers. These lasers emit 0.2 μs wide pulses of infrared light (930 nm) with a repetition rate of 40–80 Hz. Photosensitive receivers (photodiodes) are used as the detectors. The beams emitted are narrow, at 2.5 cm wide and directed within the semi-plane corresponding to the direction of travel. Obstacles can be located quite accurately because the emitted beams are quite narrow and the angular resolution of the detection system is small.

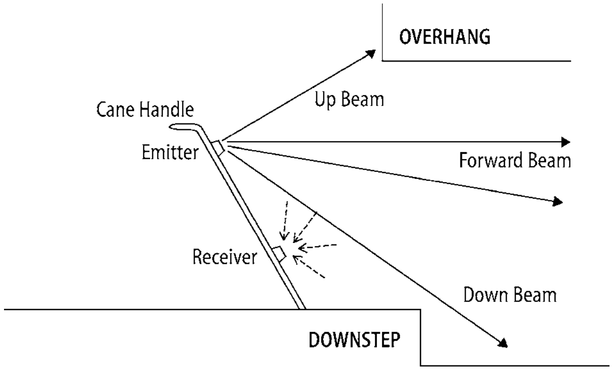


Figure 5.14. Laser cane – beam geometry

The degree of complexity of a particular obstacle detection system is related to its arrangement of detection beams and receivers. Three beam directions are often described, but different versions of the canes may carry technology for fewer beam directions. These directions are (see Figure 5.14):

- *Upward beam*. This beam is designed to detect objects at head level, typically at 1.80 m height.
- *Forward beam*. This beam detects obstacles in the travel path directly ahead. A typical setup would detect objects up to a height of 0.6 m at a distance of 1.5–3.6 m from the cane tip.
- *Downward beam*. This beam detects any down steps in the forward-path. Typically, the resolution is such that down steps of height of at least 15 cm at about 0.9 m from the cane tip can be detected.

A beam to detect obstacles at the side, as well as any narrowing of the path, would be useful, but, currently available laser canes do not have a sideways beam.

The actual detection principle is Cranberg's principle of optical triangulation. The device emits pulses of infrared light, which are detected by photodiodes located behind the receiving lens if reflected from an object in the travel path. The angle made by the diffusely reflected infrared ray passing through the receiving lens is an indication of the distance to the object. This geometry is shown in Figure 5.15.

Ultrasound cane systems

The other main approach is the of use of ultrasound, *i.e.* sound waves at frequencies greater than 20 kHz as the detection medium. Whilst optical systems are based on angular geometry, the ultrasonic system calculates the obstacle's distance from the elapsed time.

Polaroid makes one of the most frequently used ultrasonic sensors. This sensor emits a short pulse of ultrasound. If an object is located in the path of the ultrasound beam, then a portion of the beam will be reflected back to the sonar that has switched to microphone mode immediately after transmitting the ultrasound

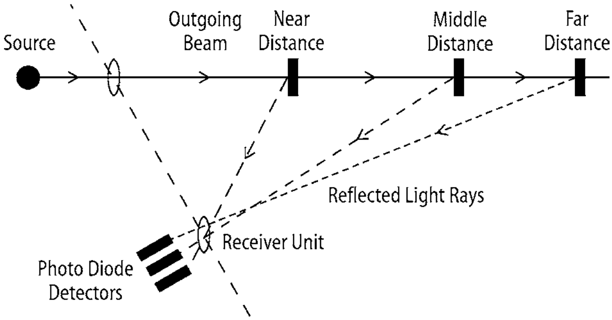


Figure 5.15. Cranberg's principle of optical triangulation

beam. When an echo is received at the sonar a measurement of the elapsed time can be used to compute the distance of the object from the emitting source. The velocity of ultrasound in air at 0 °C is constant at 331 m s^{-1} and increases at approximately 0.6 m s^{-1} per °C rise in temperature; thus the required distance is easily calculated.

Polaroid sonars have a maximum range of 10 m and an accuracy of 0.5% of the distance measured; thus at a distance of 5 m, the accuracy is $\pm$0.025 m. The beams are fairly narrow, since ultrasound waves from the Polaroid sonar propagate with a cone-shaped profile of 30° spread. Ultrasonic technology has advanced significantly in recent years and sophisticated signal processing algorithms are used in a device like the Ultracane, which is described fully in Chapter 6.

A key difference between laser and ultrasound systems is their effectiveness in detecting plate-glass and transparent plastic. Many urban environments include doors and shop windows made of plate-glass and transparent plastic is a common material for bus shelters. Transparent surfaces will generally reflect ultrasound, but transmit light. Therefore, infrared-based systems may not be able to detect and indicate the presence of a glass door or transparent bus shelter, whereas an ultrasound system will.

Information transfer in obstacle detection systems

This involves the human–technology interface through which information about obstacles or obstacle-free paths is relayed to the end-user. The interface has two important features, the *information* to be communicated to the end-user, and the *means* of communicating this information.

Both audio and tactile interfaces, as well as a combination of the two can be used to transfer information to the user. A tactile interface usually comprises vibrating buttons or pins. The audio interface usually comprises tones of different pitch, though speech and musical sounds are other possibilities. These sounds or speech should preferably be conveyed through a single earphone, so as not to impede perception of other environmental sounds and not to add to existing noise pollution.

Other important issues include how much and what types of information should be conveyed to the user. In the case of infrared and ultrasound technology canes,

Table 5.2. Detection configurations and information channels

System no.	Beam configuration	Information channels total
1	Downward, forward	2
2	Downward, forward, upward	3
3	Downward, forward, upward, reinforced upward	4
4	Downward, forward left, forward right	3
5	Downward, forward left, forward right, upward	4

the obstacle information has already been reduced to that available from a number of particular directions. In the case of some of the other technologies used to obtain information, such as the camera, the potentially available information is much more complex and decisions have to be made about how much of this information should be conveyed to the user. The tendency has generally been to focus on information about obstacle locations rather than to try to present an overview or more detailed information of the scene.

In the case of a technology cane, which obtains information from a number of different directions, a number of different information channels, which could be auditory, tactile or a combination, will generally be required. Several typical beam configurations along with the total number of information channels for each configuration are given in Table 5.2.

The number of information channels given in Table 5.2 then has to be mapped to a user interface consisting of a suitable mixture of audio and/or tactile modes. For example, for the configuration of System 2, three vibrating tactile buttons matched to the cane-bearer's thumb, and two fingers might be appropriate. Alternatively, the downward and forward channels could be tactile and the upward channel audio. Providing both tactile and audio options for all channels could have a number of benefits, including making the device available to deafblind users. However, there may be additional costs and increased complexity. Decisions on the human–technology interface are usually based on a combination of the results of end-user trials, technical constraints and cost considerations.

5.6 Other Mobility Assistive Technology Approaches

Guide dogs, long canes and technology canes are all well-established assistive technology solutions for obstacle avoidance. As indicated in Section 5.3, a number of (electronic) travel aids have been developed, though many of them have not gone beyond the prototype stage. In this section, the range of other devices that have been developed is illustrated by three examples, the Polaron, which is a clear travel path indicator, the Miniguide, which provides supplementary information to support a primary travel aid and the vOICe system, which provides a 'soundscape' of a visual scene.

5.6.1 Clear-path Indicators

The engineering response to a typical travel scene (Figure 5.3) has been to decompose the problem into a number of issues and then devise systems for these sub-problems. As discussed previously, the long cane is able to detect obstacles and identify a clear-path at lower leg height. However, other obstacles may well reach down to chest and head height as in the example of overhanging branches on trees in the environment. Laser and ultrasound technology canes are able to detect a clear-path at lower leg height and, depending on the beam configuration, at other heights too. Therefore, there is a need for assistive devices to supplement the long cane by providing information on obstacles at head and chest height. Such devices are usually termed clear-path indicators, as they are generally designed to simply indicate that the forward path at chest and head level is clear and negotiable.

The Polaron

The Polaron, manufactured by Nurion-Raycal, USA, is a clear-path indicator or an object detector that is available in a torch format or in a version that can be worn on a loose strap around the neck to be positioned at chest height. It is described here in its chest-mounted format. The echolocation principle of ultrasonic sound is used to detect whether the path is clear. In the chest-mounted format, the ultrasound transmitter and receiver are worn at chest level. The range for detection of obstacles can be set at 1.22 m, 2.44 m or 4.88 m (4, 8, or 16 ft) from the user. A sophisticated human–technology interface of audio and tactile cues is used to inform the user of obstacles in the forward path, as shown in Table 5.3.

The Polaron has sufficient tactile cues to be used by a deafblind person. The chest mounted location has the advantage of leaving the user's hands free. This enables the Polaron to be used together with a long cane or guide dog (in one hand), while still allowing the user to carry shopping or other baggage in the other. However, this location has the drawback of being conspicuous. Similarly, a blind person who uses a manual wheelchair could use the Polaron to support independent mobility.

Wheelchair mounted clear-path indicators

Whilst the Polaron can be used with a manual wheelchair, powered wheelchairs need a dedicated device. There are two main reasons for this. Powered wheelchairs can travel two to three times as fast as a pedestrian can and therefore a faster

Table 5.3. Audio-tactile interface cues – Polaron

Path status	Interface response
Clear-path: no obstacles	Polaron silent
Obstacle more than 1.83 m (6 ft) away	Polaron emits low-frequency audible sound
Obstacle between 0.91 and 1.83 m (3 ft and 6 ft) away	Polaron emits series of audible clicks and vibration at chest level
Obstacle less than 0.91 m (3 ft) away	Polaron emits high pitch bleeping and a tactile vibration occurs in the neck strap

response time is required. In addition, the additional impairments of blind and visually impaired users of power wheelchairs mean that they may not be able to respond fast enough. Therefore, 'intelligent' wheelchairs have been developed to assist users with functions such as obstacle and collision avoidance, going through a narrow doorway, between pillars or along a narrow hallway or passage as well as with landmark-based navigation. Only a small number of 'intelligent' wheelchairs are commercially available, including the Smart Wheelchair for children, whereas most of them are still in the prototype stage.

'Intelligent' wheelchairs raise the same problem of *control sharing* between the user and the chair system as do robotic walkers. They generally resolve it by giving the user control over high level functions, such as directing the wheelchair to a desired location, while providing different degrees of assistance with low-level functions such as obstacle avoidance. It should be noted that 'intelligent' powered wheelchairs have not been developed specifically for blind wheelchairs users and that they may require more assistance with obstacle avoidance and manoeuvring the chair in a narrow space than sighted wheelchair users. The 'Intelligent' wheelchair illustrates one approach to obstacle avoidance for powered wheelchair users, namely, an integrated system that forms part of the wheelchair control system. Another approach is a separate device that can be mounted on the wheelchair, for example, the Wheelchair Pathfinder as manufactured by Nurion-Raycal, USA.

Wheelchair Pathfinder

The Wheelchair Pathfinder is an obstacle detection system that can be used with manual and powered mobility vehicles and wheelchairs. It uses a mixture of infrared and ultrasonic sensing devices to give information about the path to be travelled. A downward pointing laser beam system is used for step detection and ultrasound is used to the front and sides for clear-path detection. The modes of operation are as follows:

- *Forward path obstacle detection.* This uses a forward path ultrasound beam. The forward detection distance can be set to be 1.22 m (4 ft) or 2.44 m (8 ft). An intermittent beeping sound indicates that the forward beam has detected an object.
- *Side object detection.* A continuous tone indicates the presence of an object within 30.5 cm (12 ins) to either side of the wheelchair. To distinguish between objects on the right and left side of the wheelchair user, different pitches (frequencies) are used, with the higher of the two pitches indicating an object to the right hand side. Knowledge about objects to the side can facilitate navigation in enclosed spaces or through doorways.
- *Step detection.* A downward pointing laser system is used for the detection of steps, curbs or drop-offs. Steps up to 1.22 m (4 ft) from the device can be detected and are indicated by a low pitched audio signal.

Although the obstacle detection information described above is transmitted using sounds, the Wheelchair Pathfinder is also available with a tactile interface making it suitable for those who prefer tactile information, such as deafblind people. The system can be used in several different ways, for clear-path detection and navigation.

- *Clear-path detection*. When the intermittent sound from the forward path detector is heard, then the wheelchair should be turned slowly until the intermittent beeping sounds stops. The absence of an intermittent beeping indicates that the forward path is now clear.
- *Finding a landmark*. In this case the beeping signal is used as a homing signal so that the user keeps the intermittent beeping signal in front of the wheelchair and homes in on a desired destination landmark.
- *Straight travel*. In this mode the side beam is used. The wheelchair is positioned to within 30 cm of a wall or a hallway so that the side beam issues a constant and continuous tone. Forward travel that maintains this side beam tone means that the wheelchair is travelling parallel to the wall. When the side beam tone stops, an open doorway or a corridor intersection has been reached.

5.6.2 Obstacle and Object Location Detectors

There is a clear, though sometimes subtle, difference between obstacle avoidance and object location. The primary mobility aids, such as the guide dog and the long cane, provide obstacle avoidance information. However, there are differences between this task and the object location task, where the user actually wants to make contact with an object. This could include detecting the end of a queue, the start of a sales counter or the location of lift doors, the waste paper bin or a desk chair. The use of a long cane is probably inappropriate and, in addition, could be damaged in some circumstances, for example, by closing lift doors. To gain a specification for a suitable assistive technology system, the CAT model checklist is completed and shown in Table 5.4.

This results in a specification for a small lightweight hand-held device that could be used to explore and interrogate the surroundings of the user or find objects in the user's immediate vicinity. The obvious analogy is a small torch whose light beam is used to illuminate the immediate surroundings of a sighted user.

Miniguide

The Miniguide (GDP 2006) is an assistive technology solution that meets many of the requirements of this specification. It is a small hand-held, torch-like device with dimensions of 80 mm long, 38 mm wide, and 23 mm thick. It uses the echolocation principle with ultrasonics to detect objects within the range of its beam. It is essentially a support device to be used in conjunction with a primary mobility technology, such as the long cane or a guide dog. But, it can also be used to interrogate the spatial layout of the local space around a user who might, for example, be trying to find the position of one particular item.

Table 5.4. CAT model checklist – specification for obstacle/object detection

Attribute – context	
Context – *cultural and social*	Travelling alone on public transport or on foot socially acceptable
Context – *national context*	Modern infrastructure; anti-discrimination legislation in place
Context – *local settings*	Urban community; noisy outdoor and indoor environments; stationary and moving objects; weather includes rain, snow, wind
Attribute – person	
Person – *social aspects*	Support from family and friends; training in orientation and mobility is available
Person – *attitudes*	Willing to try new assistive technology and will persevere
Person – *characteristics*	Visually impaired with tunnel vision, blind or deafblind; physically mobile. Preference for independent travel
Attribute – activity	
Activity – *mobility*	Obstacle detection; spatial awareness
Attribute – assistive technology	
AT – *activity specification – task specification*	Locates stationary objects; preferably also able to locate moving objects; able to identify some objects; obtains additional information to provide a sense of spatial awareness; identifies openings such as open doors; provides information on objects, openings *etc.* to user in an accessible and easily comprensible form
AT – *activity specification – user requirements*	Portable; small; wireless; options for sensory channels – tactile, audio; lightweight; battery lasts several hours, *i.e.* for long journey
AT – *design issues – design approach*	Design for visually impaired, blind and deafblind enduser groups
AT – *design issues – technology selection*	Options: ultrasonics, laser; battery powered. Integral battery or separate battery pack
AT – *AT system – system interfaces*	Provides environmental information, including object distances; should provide feedback using different sensory modalities; information provided should be unambiguous
AT – *AT system – technical performance*	Robust construction; weather resistant; high reliability; battery lasts several hours. *i.e.* during a long journey
AT – *enduser issues – ease of use*	Portable; wireless; fits in a pocket; easy battery replacement
AT – *enduser issues – mode of use*	Hand-held with wrist strap for safety or clips to a belt; audio *via* single earphone to enable user to hear environmental sounds
AT – *enduser issues – training requirements*	Minimal; device should be intuitive to use
AT – *enduser issues – documentation*	Available in different formats: standard and large print, Braille, on audio-cassette and on the Web

The first Miniguide was introduced to the market in 2000 and this model is illustrated in Figure 5.16a. This model was superseded in 2005 by a new realisation of the Miniguide concept that used an injection-moulded case, as shown in Figure 5.16b.

The description that follows is based on the User Guide for the new 2005 Miniguide model (GDP 2006). The Miniguide is battery powered and has four basic beam ranges for object detection: 0.5 m, 1 m, 2 m and 4 m. The torch-like body has a depressible ON/OFF switch and an ear plug socket at the back for an audio

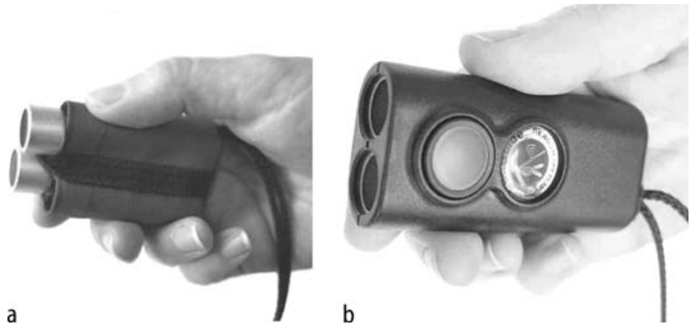


Figure 5.16a,b. Miniguide: **a** Miniguide circa 2000; **b** Miniguide circa 2005 (photographs reproduced by kind permission of GDP Research, Australia)

sound feedback signal. However, the Miniguide is basically tactile in that the torch vibrates at a speed that decreases with increasing distance of the object from the device. Slow vibrations indicate that the object detected is at the limit of the Miniguide beam range setting. A rapid vibration rate indicates that the object detected is very close to the Miniguide and no vibration indicates that there is no object within the range field of the beam. This tactile modality is suitable for visually impaired, blind and deafblind users.

However, the use of the earplug socket to obtain an audio-sound feedback signal creates an additional, or an alternative sensory information channel. Thus, the Miniguide can be configured to supply (1) only tactile feedback, (2) both tactile and audio feedback or (3) only audio feedback. The use of a single earphone allows the user to perceive the ambient background environmental noises, which provide important locational information to blind people, as well as receiving aural information from the Miniguide.

The Miniguide will be silent if it cannot detect an object in its current beam range setting. If an object is detected then an audio-feedback tone is produced. Two different types of audio tones are available:

1. *"Chirp" audio tones*. This is the default audio-tone for the Miniguide. The rate of "chirp" tones heard is the aural equivalent of the rate of vibration of the small motor within the Miniguide. A high "chirp" rate indicates that the object detected is close to the device.
2. *Sweep audio tones*. For the sweep audio setting, a continuous tone is produced, where the pitch indicates the closeness of the object detected to the Miniguide. The higher the pitch of the sweep audio tone, the closer the object detected is to the device. With practice, this variation in pitch can provide the user with some spatial awareness of their immediate environment. For example, the presence of a doorway with a wall on both sides would be indicated by the following sequence: high (wall), low (doorway), high (wall) in the audio sweep signal.

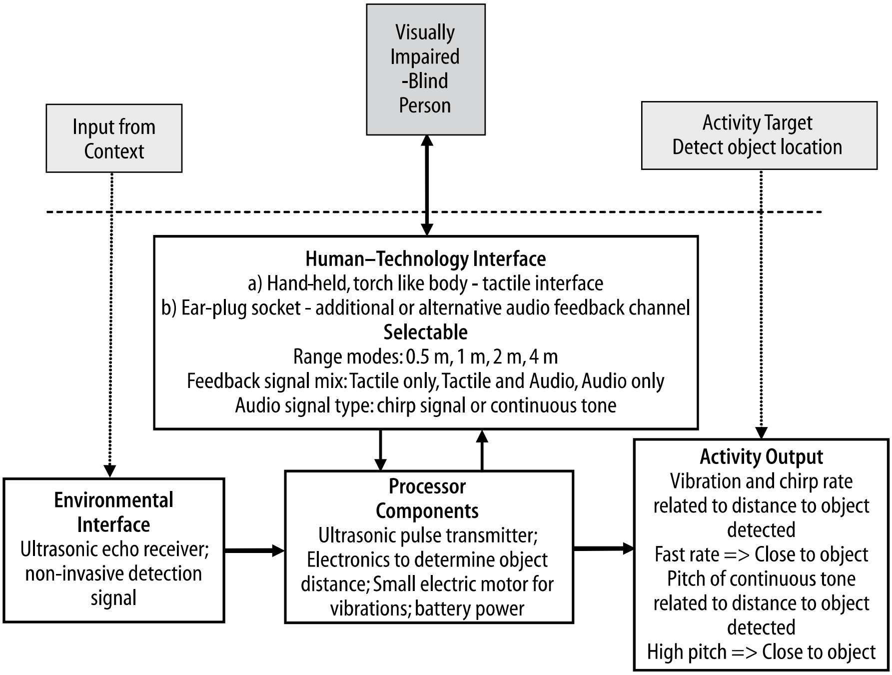


Figure 5.17. Assistive technology system block diagram – Miniguide

The 2005 Miniguide model has a cap for protecting the front of the device when it is not in use. It also has a wrist strap so that it can be released from the user's hand, without being lost or dropped, should the occasion demand it. However, its robustness is limited and damage could result if it is dropped, or heavy objects are dropped on it. Dusty conditions can cause the sensor's mesh covers to become clogged. The Miniguide is not waterproof and its use in heavy rain should be avoided. There are some (rare) conditions where the Miniguide may give incorrect interpretations. Air brakes on lorries, and car brakes can (for example) emit ultrasonic signals that might interfere with any Miniguide used nearby. However, the interference is usually very short lived and comparatively rare. The Miniguide has been designed and tested to be immune to radio frequency interference, for instance from mobile telephones and CB radios. However, interference is possible if the Miniguide is very close to a transmitter, or the transmitter is powerful.

A CAT model assistive technology system block diagram for the Miniguide device is shown in Figure 5.17.

5.6.3 The vOICe System

The vOICe system aims to map a visual image into a soundscape. It is an experimental system devised by Peter Meijer for "seeing with sound". If a fine grid (where position in the grid is given by the familiar rectangular (x, y) co-ordinates)

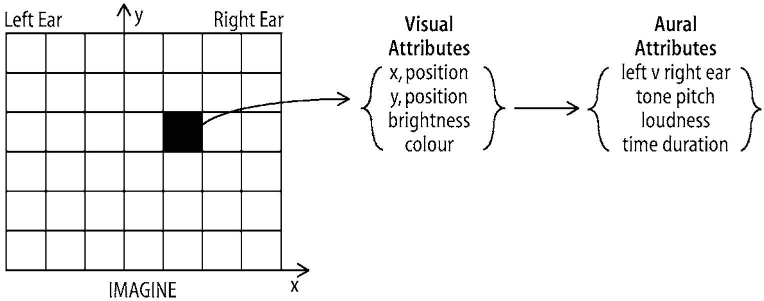


Figure 5.18. vOICe system: visual image to soundscape mapping

is placed over an image, as in Figure 5.18, then each square in the grid can be considered to have the visual attributes {(x,y) (indicating position), brightness, colour}. The principle of the vOICe system is to map each of these visual attributes to the aural attributes {left/right ear, pitch, loudness, time duration of the sound}. The summation of the sounds from the different grid squares in the image produces a soundscape version of the image.

The soundscape images can be refreshed at the rate of one image per second and the visual images can be supplied to the vOICe system from a PC camera, an image file or a computer screen. This technology raises the question of the ability of the auditory system to process and make sense of this type of very complex information. There is also the further question of whether this 'sense making' is dependent on training at an early age, since sighted people 'learn' to process complex images from birth. If the auditory sense is able to process the information, even if only with training, then this approach could be used to enhance the experience of blind people by giving them a sound picture of their environment.

5.7 Orientation Assistive Technology Systems

Designing and constructing a long-range orientation aid for visually impaired and blind people poses a significant challenge. Past projects in the US and Europe have led to the development of prototypes and currently, at least two commercial devices are available. However, this is an area where the basic enabling navigational technologies are maturing and entering mainstream consumer products. Although there are now a number of commercially available orientation devices for visually impaired and blind people, generally based on global positioning systems (GPS), published research on what the enduser community would like to have or actually need in an electronic orientation and navigation assistive technology device seems rather sparse. Clearly, long-range orientation aids need to be easily portable and, if purely devoted to providing orientation information, to leave at least one hand free to allow use of a long cane. As with obstacle avoidance systems, the device has to perform two main functions:

- Obtaining way-finding information and the relationship of the user to landmarks.
- Communicating this information to the user.

Unlike obstacle avoidance systems, (portable) assistive orientation and navigation systems are of great interest to sighted people. The existence of this very large user community has led to a number of important developments in the past few years, generally based on the use of GPS. Developments have included hand-held GPS devices, in-car navigation and map systems and position data features in mobile phones. However, GPS technology is only one option for acquiring position data and other communication technology methods are well advanced and ready for application. The popularity of this type of device has led to a reduction in prices and will undoubtedly lead to further development of the different technologies. However to exploit these technologies in an orientation and navigation system for visually impaired and blind people, an understanding of the differences in the requirements of the sighted and visually impaired communities is needed. This includes the following factors:

- *System accuracy*. Blind and visually impaired people will frequently need to know their location to a much greater degree of accuracy than sighted people do. For instance, an accuracy of one metre or less will be required to ensure that blind and visually impaired users are able to locate the door of the house they are looking for.
- *Information frequency and quantity*. Blind and visually impaired people often use a greater number and variety of landmarks than sighted people when on a journey. It is not likely that all the most appropriate landmarks, such as pedestrian crossing location data, will be available from the digital maps so often used in these systems.
- *Spatial orientation resolution*. Blind and visually impaired people may require information on the direction they are facing. This type of information was not provided by the earlier GPS systems. This problem can be resolved by the user travelling a short distance and being informed whether the distance to their destination has increased or decreased. However this is not a particularly user-friendly solution.
- *Unambiguous audio and tactile information interfaces*. Many visually impaired people will not be able to use the visual information often provided by these systems, particularly while travelling and will therefore need all the information in verbal or tactile form. This will require the verbal information and descriptions to be presented in a particularly clear and unambiguous way, as this is the sole source of information and ambiguities cannot be resolved by looking at a visual route map.
- *Back-up system provision*. Blind and visually impaired people are likely to experience problems that are more serious than those experienced by sighted people if the electronic orientation system fails. For example, the GPS system can fail in built-up areas, due, for instance, to the line of sight to the satellites being

blocked by high buildings. Therefore electronic orientation systems designed for blind and visually impaired people will require a backup system.

5.7.1 Global Positioning System Orientation Technology

GPS systems use information from at least three satellites to detect the user's location. The satellite system is interfaced with a digital map that can then relate the detected location to appropriate landmarks. In-car systems incorporate speech delivered information, as well as visual information to try to avoid the driver being distracted from concentrating on the road. Therefore, systems of this type are potentially equally relevant to visually impaired, blind and sighted people and some GPS devices are available.

The satellite system currently used in GPS is a military system belonging to the US Department of Defense. This system was originally available to civilians at a reduced accuracy, called selective availability, of ± 100 m rather than ± 5 m. Selective availability was removed in May 2000, enabling all users to benefit from the full accuracy of the system. Accuracy can be increased by the use of differentially corrected GPS (DGPS) which corrects bias errors at one location using measured bias errors at another location and also eliminates static errors from satellite signals entering the earth's atmosphere. The DGPS services are provided commercially with different services giving accuracies of ± 1 m to ± 4 m, with corresponding prices. Therefore, in principle, DGPS is able to provide sufficient accuracy to be used by blind and visually impaired people. However, there will still be a need for modification of existing digital maps to provide sufficient and appropriate information for blind and visually impaired people. There is also a need for further development of the information interface to provide sufficient and appropriate information to enable the user to navigate to their destination safely while being aware of any landmarks of interest on the route.

Braille Notetaker with GPS

This device has the advantage of being useable by deafblind as well as blind people. Chapter 8 contains a more detailed presentation of the development history for this device.

5.7.2 Other Technology Options for Orientation Systems

Mobile telecommunication technology orientation systems

Another potential approach to obtaining position data is the use of mobile phone signals in combination with the technical development resulting from legislation to ensure that the emergency services are able to pinpoint the location of calls made from a mobile telephone. As with GPS, the resulting information has to be interfaced with a digital map in order to relate the location to appropriate landmarks. In addition, as in the GPS case, it is not always possible to obtain the location using mobile phone signals. Combining the two approaches might prove a solution. Alternatively, a backup system would be required.

Talking compass

The C2 Talking Compass from Robotron Group(Australia) is a hand-held compass that uses digitised speech to state the direction in which the compass is pointing. It is button-activated and can give the eight major compass points (N, NE, E, SE, S, SW, W, and NW) in two languages selected from most of the major languages of the world. More details of compasses can be found in Chapter 18, Section 18.4.4.6.3.

Embedded systems in the environment

Quite a different approach is to use a system of signal beacons embedded in the environment. There are two architectures possible, one where the traveller uses a hand-held receiver to interrogate the beacon and a second, where the beacon itself delivers an audio message when activated by a traveller. The two approaches are compatible and a combination of embedded beacons plus a hand-held receiving device could considerably enhance the available information and journey experience for visually impaired and blind travellers. In addition, embedded orientation devices, particularly in urban areas, could be considered part of the design of accessible environments. There is increasing awareness of the importance of accessible environments and some progress has been made, though considerable further work will be required. An overview of developments in environmental accessibility can be found in Section 5.8 and more details are given in Chapters 10 and 11.

5.8 Accessible Environments

The earliest legislation on accessibility for blind people may be that found in the Bible (Lev. ch 19, ver 14) prohibiting leaving 'stumbling blocks' which could trip up blind people. Although in many ways considerable progress has been made since then, potholes and badly maintained road surfaces can still pose serious problems to sighted as well as blind and visually impaired people.

The previous sections of this chapter have reviewed the assistive technology systems that can be used to overcome some of the barriers presented by inaccessible environments. The development of the social model of disability and increasing activism by organisations of disabled people has lead to increasing, though still patchy awareness of the importance of full civil rights and social inclusion for disabled people. Many countries have legislation prohibiting discrimination based on disability, though much of this legislation is also criticised by groups of disabled people and implementation is variable. Being able to access environmental facilities and services is an important aspect of social inclusion.

Many countries now have legislation prohibiting discrimination on the grounds of disability. This generally includes legislation on the accessibility of the facilities, services and physical environments of cities, towns and villages, as well as the associated public infrastructure. In some cases, accessibility requirements have also been added to building regulations. Examples include the Accessibility Guidelines generated by the U.S. Access Board in response to the Americans with Disabilities Act (1990) and the Architectural Barriers Act (1968). Within these guidelines are regulations to require buildings, including offices, courtrooms, prison cells and sports centres, to be made accessible to visually impaired and blind people. Key aspects are building layout, lighting, stairs, lifts, escalators, doors, doorways, and signage. Improvements in these areas are likely to have benefits for everyone, whether blind, visually impaired or sighted.

As indicated at the start of this section, an important aspect of accessibility is ensuring that the urban environment, including pavements, roads and buildings, is in good repair, so that there are no unintended hazards. This is clearly of benefit to everyone, regardless of whether they are blind or sighted. Other aspects of accessibility from which everyone can benefit include the improved positioning of street furniture, such as lamp posts, traffic lights, benches and waste bins, to reduce the likelihood of collisions. Many countries, which experience cold weather, with the possibility of ice and snow in winter, spread grit on the roads and use snowploughs to keep the roads clear, butt end to ignore the pavements. Pavements covered with ice and snow are a danger to all pedestrians and to blind and elderly pedestrians in particular. However, there seems to be limited awareness that keeping pavements clear should be considered part of accessibility.

Good street lighting at night is an important part of accessibility, particularly as many visually impaired people who are reasonably confident during the day may feel reluctant to go out without a guide at night. The development of appropriate lighting systems will require the involvement of end-users and taking into account the differing needs of different groups of people. Lighting should probably be diffuse rather than concentrated to avoid causing problems and potential inaccessibility to people who are light-sensitive.

Another important aspect of accessibility is urban design. Unfortunately, there has been a tendency to site facilities, such as shopping centres out of town, often on routes that are often poorly served by public transport. This has been rightly criticised on environmental grounds, but also has important implications for accessibility. Compact urban design with facilities located close to centres of population and a consequent reduction in the need for travel to access facilities and services will benefit disabled people as well as the environment. One advantage is the reduction in the distances to be travelled and the enhanced accessibility for those travelling on foot or by public transport. A further very important benefit is the reduced need to learn different and possibly very complex routes. It is much easier to gain familiarity with one relatively compact area and this is likely to increase the confidence of blind and visually impaired people in going out on their own. In addition, more compact urban design may lead to a resurgence of communities. This is also likely to be of benefit to blind and other disabled people in reducing isolation and increasing the likelihood of finding people they know in the event of getting lost or needing assistance.

Many urban environments have very poor signage. Street names are often missing, defaced or blocked. They tend to have similar locations within a given town, but locations vary in different towns. Signs on and within public buildings are often very small, ambiguous or absent. The use of clearer, larger street signs and signs on and within public buildings would improve accessibility for many visual impaired people, as well as having benefits for sighted people.

While the focus of this book is blind and visually impaired people, environmental accessibility is for everyone. In some cases, developments that benefit blind and visually impaired people are of benefit to the whole population. In other cases, potential conflicts between the needs of different groups of people should be resolved. For instance, the increasing use of audible indicators of train doors being open and time to cross at pedestrian crossings is a welcome development to many blind and visually impaired people. However, it can cause problems for people who are noise sensitive. A solution, which would be of also be of benefit to deafblind people, would be for environmental objects, such as train doors, lifts and pedestrian crossings, to be fitted with an infrared transmitter. Appropriate standardisation would ensure that blind and visually impaired people would require only one receiver to receive all the different output signals. The user would then have a choice of vibro-tactile or auditory output through a single headphone (so as not to obscure other environmental sounds). An increase in frequency of the vibration or auditory sounds could be used to indicate that the 'on' or 'open' or 'go' phase was about to end. Tuning of the transmitting light emitting diodes generating the infrared signal could be used to give a short range. This would ensure that users only received a signal from the appropriate door or crossing.

The above suggestions include some of the many ways in which urban and other environments can be made more accessible to blind and visually impaired people through improved design and without the need for assistive technology. However, assistive technology also has an important role in making the physical environment accessible and should be seen as complementing other approaches, of the type discussed here.

Assistive systems to increase environmental accessibility for blind and visual impaired people can be divided into the following two main categories, that are discussed in more detail in Chapters 10 and 11:

1. Assistive systems within the physical environment
2. Embedded navigation and information assistive systems

Assistive systems within the physical environment

In considering accessibility of the urban environment, it is useful to consider the accessibility of different aspects of the environment separately. This section will consider the following:

- Streets.
- Buildings, including their entrances, exits and interior layout.

Accessibility for blind and visually impaired people to the street environment can be improved by the use of tactilely marked surfaces. For instance, strips of tactile tiles can be used to indicate pedestrian crossings. In Japan in particular, a continuous strip of tactile tiles is being used to indicate a safe route along a pavement. The use of audible and tactile indicators at pedestrian crossings is important for visually impaired and blind pedestrians. Some aspects of building accessibility have already been discussed. There is also a role for navigational cues within large buildings to indicate main corridors, departments or sections and important rooms.

Embedded navigation and information assistive systems

Long-range navigation is as important as obstacle avoidance and can be even more complex. However, it is only relatively recently that there has been some success in the development of orientation systems, whether portable or embedded. One approach is to embed a set of navigational beacons in the environment that can be activated to provide:

- Local cues for doorways, crossings and transport entry and exit points, such as bus stops and metro stations.
- Long-range information about directions to locations such as banks, shops, public buildings and information about public transport at that location.

Whenever feasible, the user interface would be provided through a hand-held device with both tactile output and audio output through one earphone. Standardisation would allow one device to receive information from all the different navigational beacons. Multiple channels could allow users to switch between local and long-range information or between local cues and traffic crossing or lift opening alerts.

An example of this type of beacon is given by talking signs, which provides a repeating, directionally selective voice message, which originates at the sign and is transmitted by infrared light to a hand-held receiver some distance away. The directional selectivity is a characteristic of the infrared message beam where the intensity and clarity of the message increases as the receiver points at the sign more accurately. The receiver is light and small, the sign is easy to install, consumes little power, and is easy to program with human voice or synthesised voice messages. Talking signs uses light-emitting diodes to transmit digitally encoded human speech messages that are intercepted and then relayed through a speaker in the hand-held receiver. The hand-held receiver contains a photo-detector at its front end so that the message is detected when the receiver is pointed in the direction of the sign transmitter. The transmitter and LED arrays can be tuned to control the maximum distance of reception of the message and the direction(s) of transmission. Only users with a receiver can access the stored message and the system is illustrated in Figure 5.19.

There are other approaches to embedded information systems, for instance based on Bluetooth. A case study of the use of Bluetooth to increase the accessibility of public transport is described in Chapter 11.

These embedded navigation and information systems are systems that are *distributed* throughout a physical environment. However, a Talking Tactile Map (Landau 1999) can be used to give more localised information. The user touches the tactile map and obtains audio information about the location at the point of contact. Route information can be provided for both outdoor and indoor environ-

Figure 5.19. Talking Signs® in action (photograph reproduced by kind permission of Talking Signs, USA)

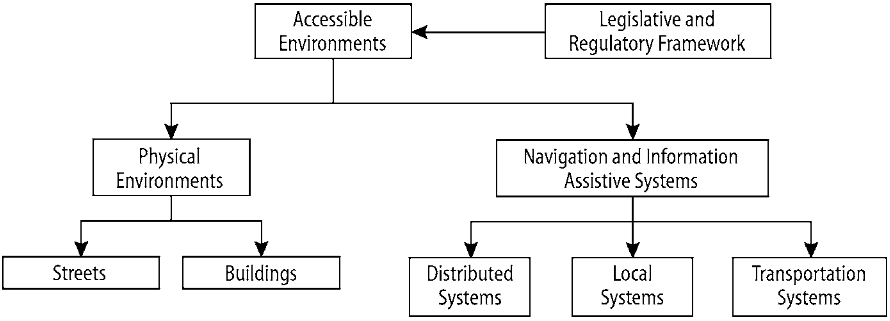


Figure 5.20. Accessible environments for visually impaired and blind people

ments. For instance, outdoor information could include details of squares in the city, whereas the indoor version would provide information on the layout of offices which members of the public are likely to want to access.

Although some progress has been made in making streets, buildings and public transport systems fully accessible, considerable further work is still required. In some cases, this will involve the rigorous implementation of existing regulations and in other cases, it may require further technical developments and extensive consultation with disabled people.

The areas of accessibility, which are of particular importance to visually impaired and blind people, are illustrated in Figure 5.20 and further detailed discussions of the associated assistive technologies are given in Chapters 10 and 11.

5.9 Chapter Summary

In the past, an *ad hoc* approach has been used to provide engineering and technological solutions to support (independent) travel by visually impaired and blind people. The travel scenario was decomposed and various devices constructed for different aspects of travel, without serious consideration being given to a co-ordinated approach to provide an integrated set of technological solutions. As with the survival of the fittest, the currently available devices (commercial or otherwise) tend to be those that have been taken up enthusiastically by the end-user community of blind and visually impaired people as genuinely meeting their travel needs. For example, the guide dog, the long cane and, to a certain extent, technological canes such as the Ultracane and the hand-held Miniguide (a secondary device) are successful obstacle avoidance assistive technology. However, guide dogs and the long cane, rather than technological canes, are still the main obstacle avoidance devices used by blind and visually impaired people. The main principles of these and some other assistive technology systems for obstacle avoidance have been described in this overview chapter. Some of the technological solutions will be considered in more detail in subsequent chapters of the book.

Orientation and navigation systems for blind and visually impaired people are required to provide the user with accurate information on their location. Systems for providing this information, based on the global positioning system or the signal strength in mobile telephone networks have been described in the chapter. However, currently the only commercially available orientation systems for blind and visually impaired people are based on GPS technology. Although these devices have been developed as stand-alone applications, there would be advantages to end-users in them being designed as one component of an integrated location and information system for an urban area within an accessible environment for visually impaired and blind people.

Long-range travel involves considerable information processing activities to enable travellers to reach their desired destinations. The fundamental psychological studies described in this chapter have shown that blind people with no prior visual experience are more likely to use body-centred co-ordinates to develop a mental map of a travel environment, as part of the process of successfully traversing it. Consequently, it was suggested that the journey mapping skills of visually impaired and blind people could be enhanced by training with tactile maps.

The importance of non-visual external reference points in the environment for the travel safety of visually impaired and blind people leads naturally to the concept of the accessible environment. This requires modifications of the urban environment, including the provision of high density tactile and non-visual cue systems to enable visually impaired and blind people to understand the terrain through which they are passing in order to negotiate it more easily. The development of accessible environments will also require the provision of appropriate locational and infrastructural information to facilitate independent mobility and navigation. The overview presented in this chapter has identified environmental cue systems, navigation (beacon) systems, information systems and transportation information systems as key components of the accessible environment for visually impaired and blind people.

Finally, although the accessible environment is an exciting social and technological development, it is still in its early stages. Continued progress will require involvement of blind and visually impaired people to ensure that developments meet their real needs. There is also a need for research investigations, involving blind and visually impaired people, on demonstration systems. Accessible environments are discussed in more detail in Chapters 10 and 11.

Acknowledgement. The authors would like to acknowledge the kind assistance and permission of Professor Johann Borenstein (University of Michigan, Ann Arbor, USA) to reproduce the two figures (Figures 5.8 and 5.9) on the Guidecane development. They would also like to thank GDP Research, Australia for permission to reproduce the photographs of the two versions of the Miniguide device.

Questions

Q.1 Draw the detailed block diagram for the assistive technology component of the CAT model.
Q.2 Summarise the differences between body-centric and external reference systems for mobility. Which is the more robust and why?
Q.3 List three advantages and three disadvantages of the use of a guide dog as an obstacle avoidance assistive technology.
Q.4 Compare and contrast the physical principles involved in using infrared light and ultrasound in a cane-based obstacle detection system.
Q.5 Take a typical street scene and list all the potential hazards and barriers to safe mobility for a visually impaired or blind person.
Q.6 List and describe briefly the main components of environmental accessibility.

Projects

P.1 Use the Internet to construct a historical chronology of the development of obstacle avoidance devices and aids. Correlate the sequence to a historical chronology of the technological developments used.
P.2 Obtain information on a number of different types of devices that use GPS or mobile telephone technology for giving positioning data. Determine which of these devices are accessible to:
   (a) visually impaired people
   (b) blind people
   For devices that are not accessible, consider how they might be modified for blind and visually impaired travellers.
P.3 Conduct an accessibility audit of your local library for:
   (a) visually impaired users
   (b) blind users
   Make recommendations for resolving any accessibility problems identified by your audit.

References

Aigner, P., and McCarragher, B., 1999, Shared control framework, Applied robotic aid for the blind, IEEE Control Systems Magazine, Vol. 19, No. 2, pp. 40–46, April

Benjamin, J., 1968, A Review of the Veterans Administration blind guidance device project, Bulletin of Prosthetics Research, Vol. 10, No. 9, pp. 64

Benjamin, J.M., Ali, N.A., and Schepis, A.F., 1973, A laser cane for the blind, Proc. San Diego Biomedical Symposium, Vol. 12, pp. 53–57

BFB, 2006, wwwbatforblind.co.nz; website address for Bay Advanced Technologies Ltd, 31/7, Saint Vincent Avenue, Remuera, Auckland 1005, NZ

Blasch, B.B., Long, R.G., and Griffin-Shirley, N., 1989a, National Evaluation of Electronic Travel Aids for Blind and Visually impaired individuals: Implications for Design, RESNA 12th Annual Conference, New Orleans, Louisiana, pp. 133–134

Blasch, B.B., Long, R.G., and Griffin-Shirley, N., 1989b, Results of a national survey of electronic travel aid use, J. Visual Impairment and Blindness, Vol. 83, pp. 449–453

Borenstein, J., and Ulrich, I., 1997, The GuideCane, a computerised travel aid for the active guidance of blind pedestrians, IEEE Conf. Robotics and Automation, Albuquerque, New Mexico, April

Brambring, M., 1985, Mobility and orientation processes of the blind, pp 493-508, In Warren D.H. and Strelow E.R. (Editors), Electronic spatial sensing for the blind, Dordrecht, The Netherlands, Nijhoff

BRY, 2006, www.brytech.com; website address for Brytech, 600, Peter Morand Crescent, Suite 240, Ottawa, ON, KIG 5Z3, Canada

Farcy, R., Leroux, R., Jucha, A., Damaschini, R., Gregoire, C., and Zogaghi, A., 2006, Electronic travel aids and electronic orientation aids for blind people: Technical, rehabilitation and everyday points of view, Keynote Paper, Proceedings of Conference on Assistive Technologies for People with Vision and Hearing Impairments, 18–21 July, 2006, Kufstein, Austria

Farmer, L.W., and Smith, D.L., 1997, Adaptive technology, In B.B. Blasch, W.R. Wiener, R.L. Welsh (Eds.), Foundations of Orientation and Mobility, 2nd Edn., pp. 231–259

GD, 2006, www.guidedogs.com.au, website for Guide Dogs NSW, Head Office, 2–4, Thomas Street, Chatswood, NSW 2067, Australia

GDP, 2006, www.gdp-research.com.au

Harper, S., and Green, P., 2002, A travel flow and mobility framework for visually impaired travellers, University of Manchester, Manchester, UK

Heyes, A., 1993, Sonic Pathfinder Training Manual, Royal Guide Dog Association of Australia, Kew, Australia

Jacobson, W., 1979, Complementary travel aids for the blind person: The Sonic guide used with a dog guide, J. Visual Impairment and Blindness, Vol. 73, No. 1, pp. 10–12.

Kay, L., 1974, A sonar aid to enhance spatial perception of the blind: Engineering design and evaluation, Radio and Electronic Engineer, Vol. 44, No. 11, pp. 605–627

Kay, L., 1980, The Sonic guide, long cane and dog guide: Their compatibility, J. Visual Impairment and Blindness, Vol. 75, No. 7, pp. 277–280

Landau, S., 1999, Tactile graphics and strategies for non-visual seeing, Thresholds, Vol. 19, pp. 78–82, MIT School of Architecture, Cambridge, Mass., USA

Li, S., 2001, Undergraduate Project Report, Department of Electronics and Electrical Engineering, University of Strathclyde, Glasgow, Scotland, UK

Mims F.M. III, 1972, Eyeglass mounted mobility aid, J. American Optometric Association, June, pp. 673–676

NR, 2006, www.nurion .net, website of Nurion-Raycal, Station Square Building 2, Suite B, Paoli, PA 1930I, USA

Russell, L., 1966, Travel Pathsounder and evaluation, In R. Dutton (Ed.), Procs. Conference on the Evaluation of Sensory Devices for the Blind, pp. 293-297, St. Dunstan's, London

UC, 2006, UltraCane™, www soundforesight.co.uk; website address for Sound Foresight Ltd, Douris House, Roecliffe Business Centre, Roecliffe, YO51 9NE, UK

Ulrich, I., and Borenstein, J., 2001, The GuideCane-applying mobile robot technologies to assist the visually impaired, IEEE Transactions on Systems, Man, and Cybernetics, Part A Vol. 31, No. 2, pp. 131–136

Ungar, S. 2004, Cognitive mapping without visual experience, In Kitchin, R., and Freundschuh, S., (Editors), Cognitive Mapping : Past, Present and Future, Routledge, London, UK

6 Mobility AT: The Batcane (UltraCane)

Learning Objectives

The long cane is possibly the most important assistive technology aid for obstacle avoidance used by visually impaired and blind people. Its principles are simple, it is low maintenance and it is effective. However, sweeping a path through urban and city pavement areas using swings of the long cane can be socially divisive. The UltraCane™ development described in this chapter is designed to retain all the good qualities of the long cane, eliminate the social disadvantages and yet deliver the same, if not more, information about obstacles in the environment to the user. The chapter is in the form of a case-study of the application of advanced technology to develop a new assistive technology system and the learning objectives for the chapter are:

- Gaining an understanding of the principles of ultrasonics as required for the Ultracane device.
- Learning about sonar guidance as used by bats and how these concepts can be used for assistive technology obstacle avoidance systems.
- Obtaining an appreciation of how the design, development and construction stages are progressed for a new assistive technology device as exemplified by the Ultracane development.
- Appreciating the issues involved in end-user trials of a prototype device and in the ultimate commercialisation stages of the development process.

6.1 Mobility Background and Introduction

The aim of this project was to create an aid for the visually impaired that would harness a range of technology opportunities to deliver a step forward in assistive technology. The original concept was devised by a multidisciplinary team having a wide ranging and complementary set of expertise in systems engineering; ultrasound technology, biology, and aural physiology.

Perhaps the most commonly used mobility aid is the long cane. The team recognised that this primary aid would form an effective and familiar basis for the introduction of new technology to create an electronic guide cane. This could augment the simple features of the long cane to increase its effectiveness as a primary mobility aid and also to enhance its safety. The first key requirement was to reinforce the *envelope of safety* by including a detector for overhead obstacles ahead of and in front of the user. The second key requirement was to extend the effective range of observation of the electronic guide cane to allow its user to move more freely, and where safe to do so, without the need to sweep the cane in a large arc. The design concept was to offer the user prioritised information concerning obstacles in their field of interest.

A self-imposed limiting requirement was that an electronic cane must not monopolise the user's hearing faculties. In order to deliver information from an electronic guide cane equipped with ultrasonic sensors to its user, without the need to commandeer hearing faculties, a tactile user interface was selected for development.

A major inspiration for the work was knowledge of the impressive capabilities of echolocating bats, which rely upon ultrasonic energy to sense their surroundings. This form of spatial interrogation was selected at an early stage as having ideal characteristics for this purpose.

6.2 Principles of Ultrasonics

Ultrasound in simple terms is a sound wave whose frequency higher is than the maximum range detectable by human beings, typically 15–20 kHz for adults. Sound waves are known to reflect from a wide variety of everyday objects. Ultrasound waves have the obvious practical advantage that they may be used without disturbing other people, or indeed animals at the frequencies to be used. Ultrasound is commonly used in medical imaging and this use has defined the basic principles of such systems (Wells 1969). Although the fundamental science is applicable the use of ultrasound in air imposes major constraints.

6.2.1 Ultrasonic Waves

An ultrasonic beam in air advances as a longitudinal wave front, in common with all sound waves. Its use to sense an object in the path of a cane user will depend upon a detectable interaction. A variety of complex scattering interactions may arise depending upon the exact circumstances. If an ultrasonic beam is transmitted from a guide cane the simplest option is to detect any resulting reflection received from the corresponding direction. A *sonar* form of this requirement can be realised by transmitting a short pulse of ultrasound energy and measuring the time, t, taken for an echo to be received. The time measured corresponds to the outward and return paths. If the velocity of the sound, c, is known then the distance, l, to the object can be calculated:

$$l = \frac{ct}{2} \tag{6.1}$$

Sound travels more quickly in dense stiff materials, and hence travels relatively slowly in gases such as air. Density is clearly related to temperature. The velocity of sound in air (in $\mathrm{m\,s^{-1}}$) is modelled by the empirical relation with temperature, T (in °C):

$$c = 331.4 + 0.6T \tag{6.2}$$

At 10 °C the velocity is thus approximately 337 $\mathrm{m\,s^{-1}}$, corresponding to a travel time of 3 $\mathrm{ms\,m^{-1}}$. To resolve a distance of 1 cm an internal electronic pulse detection system must therefore be capable of resolving times to the order of 30 µs, a relatively straightforward engineering specification.

The variation in velocity with typical environmental temperature changes implied by Equation 6.2 has a secondary effect. If the temperature of operation is assumed to be 10 °C, and the corresponding value of velocity of sound is used internally for the computation of Equation 6.1, an error can be expected if the actual temperature is above or below this design value. At temperature values of 20 (±10) °C the maximum error is of the order of 2%. This is an acceptable value, for example for a distance of 5 m at temperatures of 0 °C and 20 °C respectively the maximum error is ±9 cm.

A standard fundamental relation links the velocity of the wave to its frequency, f, and its wavelength, λ:

$$f = c/\lambda \ . \tag{6.3}$$

A typical frequency for ultrasound used in air is 50 kHz, giving a wavelength at 10 °C of approximately 7 mm.

6.2.2 Attenuation and Reflection Interactions

A critical consideration in the design of any ultrasonic sensing system is the power required to generate a detectable echo under the most demanding specified conditions. This is particularly critical in portable devices where it is likely to be a major determinant of battery life. This has two major factors.

The first is concerned with the reflection from a target object. Interfaces due to objects which are much smaller than the acoustic wavelength interact according to Rayleigh scattering; basically related to the fourth power of frequency. Where the wavelength is small relative to the smallest object particles of interest, straight-line *ray* behaviour can be assumed. An ultrasound beam may be expected to be incident at interfaces in a non-normal fashion and will split into reflected and transmitted components. As in the analogous optical case, for the reflected wave, the angle of reflection will equal the angle of incidence. The transmitted wave will obey Snell's Law relating the angles of incidence and refraction to the velocity of sound in the two materials. In a typical liquid or solid this could be several times greater than its velocity in air, *e.g.* 1500 $\mathrm{m\,s^{-1}}$ in water. In general terms the greater the difference in velocity of sound at the interface, the greater will be the amount of energy reflected. Conversely where impedances are similar most of the energy will be transmitted.

Hence ultrasound energy in air will be reflected to a high degree by solid surfaces such as metals and structural materials, and to a lesser degree by softer objects such as furnishings and clothes. The surface shape and incidence angle are also important. An ideal reflection will arise from a beam which is normal to hard polished surface. Small hard objects such as cylindrical signposts will reflect well but much of the reflected energy will be scattered away from the transmitter due to the curved surface. Hence the interaction is complex and also depends on the physical size and shape of the interface and upon the angle of incidence of the ultrasonic wave front.

The second factor that impinges upon the required transmitted power concerns the attenuation of the ultrasonic energy in the medium, in this case air. This may be modelled by Lambert's exponential law of absorption that relates transmitted to received energy *via* a homogeneous path. The rate of the exponential signal decay is a function of the medium and the frequency of the ultrasound energy. Sound waves are transmitted well by stiff materials, through which the vibration is conducted easily. Gases are relatively poor conductors of sound waves which are progressively attenuated as they travel. Higher frequencies are attenuated more rapidly. The ultrasonic frequency band from 30 to 100 kHz is found to be the most useful in terms of tolerable attenuation for applications in air.

6.2.3 Transducer Geometry

In principle the wavelength of the ultrasonic energy is directly linked to its resolving power. Thus the longitudinal wavelength is linked to the corresponding *axial* resolution. However, in this application the interest is in locating hazards rather than imaging their details. As discussed above the key point is to gain a reflection and to resolve its distance through the measured time of its echo.

The *lateral* resolution will also be of interest and demands consideration of a two-dimensional field model which encompasses the profile of the beam. A reasonable approximation is of a disc radiator whose beam width is inversely related to its diameter. Hence a point source is an omnidirectional radiator. The disc size can be used to form the desired beam width. As an approximation, the angle of the main beam to the axis of radiation can be derived from the following equation:

$$\alpha = \sin^{-1} \frac{c}{df} \tag{6.4}$$

where α is the main beam angle with respect to the angle of propagation, c is the velocity in $\mathrm{m\,s^{-1}}$, d is the diameter of the transducer in meters, and f is the frequency of sound in Hertz.

The full beam angle is therefore 2α. Thus, a 25 mm diameter transducer operating at 40 kHz produces a main beam angle of 40°. This equation is valid for situations where the wavelength is smaller than the diameter of the transducer.

Although this simple model offers a primary design insight the actual structure of the beam is much more complex and requires secondary consideration. Where the wavelength of sound is much smaller than the diameter of the transducer side-lobes are produced, radiating energy off-axis. The energy level also decays with angle, hence there is no hard cut-off at a specific beam angle. Beam angle calculations are thus approximations to ideal requirements. The transducer must of course be attached in a mounting that is likely to add further secondary effects. Often the only way to establish angular spread is to measure it. More detailed descriptions on the calculation of beam patterns can be found in Rossing and Fletcher (2004) derived from equations in Morse (1948).

Selection of the required beam width will thus be a design compromise. It must be wide enough to detect objects in the likely path of the user, without including too much of the wider field, in particular since this may feature good reflectors which do not actually pose a hazard and would result in a false positive alert. The reciprocal nature of the beam patterns means that those calculated for the transmitter also apply to the receiver. In pulse-echo systems where a single transducer is used as both transmitter (speaker) and receiver (microphone), the beam patterns are identical.

6.3 Bats and Signal Processing

The fact that bats use ultrasound to detect objects and their insect prey was first described by Donald Griffin and Robert Galambos in the 1940s (Griffin 1958). At that time, the only microphone capable of recording signals in the ultrasonic range was one that used the piezoelectric characteristics of the Rochelle salt crystal. Using these simple microphones to monitor the acoustic emissions of bats, Griffin discovered that they could detect and avoid obstacles, such as wires, by listening to the return of echoes from these objects.

Broadly speaking, the ultrasound system works in the same way as an active sonar system. The bat emits an ultrasonic pulse and times how long it takes for the echo to return. By its implicit knowledge of the velocity of sound in air, the bat is able to calculate the distance to the object.

6.3.1 Principles of Bat Sonar

The signal structures produced by bats vary considerably between species. The reasons for the differences in fine structure are still ambiguous. Most observed signals fall into the ultrasonic range, with the highest frequency used being 212 kHz, produced by the Short-eared Trident Bat (*Cloeotis percivali*), although calls made by other species can be as low as 10 kHz.

The calls have a varying signal envelope. A convenient way of measuring signal amplitudes in these circumstances is the peak-equivalent sound pressure level (peSPL). This is realised through a constant amplitude sine-wave matched to the maximum amplitude of the signal under study. The RMS amplitude of this matched signal is then measured using a conventional sound pressure meter.

The intensity of calls is generally around 110 dB peSPL at 10 cm (Waters and Jones 1995), and for most species are of the order 100–110 dB peSPL. These very high ultrasound output levels are required of the bat due to the effects of atmospheric attenuation and the low *target strength* properties of the insect under observation. In order to model the effects of these parameters, a common form of the sonar equation can be used:

$$SNL = SL - 2TL + TS - (NL - DI) \tag{6.5}$$

where SNL is the signal-to-noise ratio of the returning echo, SL is the source level, $2TL$ are the two-way transmission losses, TS is the target strength, NL is the noise level and DI is the directivity index. All parameters are in decibels.

The two-way transmission losses are a function of both spherical spreading, and of atmospheric attenuation. The mouth of the bat effectively acts as a point source in the far field, so the sound amplitude is reduced by 6.02 dB for every doubling of distance to the target. If the bat calls at 110 dB peSPL at 10 cm, then at 20 cm this is reduced to 104 dB peSPL. Levels at arbitrary distances can be calculated from the following equation:

$$SL_{\mathrm{D2}} = SL_{\mathrm{D1}} - 20 \times \log_{10} \frac{D2}{D1} \tag{6.6}$$

where SL_{D2} is the sound level at distance $D2$, and SL_{D1} is the sound level at distance $D1$. For two-way loss where the echolocation call returns from a target, the target itself acts as another point source. This results in a two-way transmission loss of 12.04 dB *per* doubled distance which can be simulated by substituting the value of 40 for 20 in Equation 6.6, with $D2$ becoming the target distance.

Atmospheric attenuation increases roughly exponentially as a function of frequency as illustrated in Figure 6.1. At 20 kHz a propagated sound at 20 °C, 70% relative humidity and normal atmospheric pressure, would lose 0.43 dB m^{-1} in addition to losses due to spherical spreading. At 100 kHz the loss increases to 3.8 dB m^{-1}, and at 200 kHz to 9 dB m^{-1}. This is the one-way attenuation, so as the echo returns to the bat, a further loss of the same magnitude occurs. For a bat using a 100-kHz echolocation signal, this represents a loss of 7.6 dB m^{-1} over the two-way spherical spreading loss.

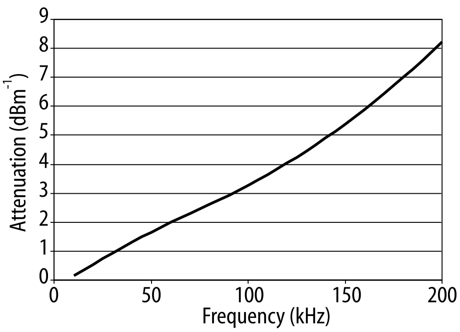


Figure 6.1. Atmospheric attenuation at 20 °C, 50% relative humidity and standard atmospheric pressure

The next parameter affecting the strength of the returning echo is the *target strength*. This is represented as the loss in sound energy between the incident sound pulse and its reflected value. The target strength is referenced to a standard distance (usually 1 m). For example, if the incident sound intensity at a target is 80 dB SPL, and the intensity of the echo recorded at 1 m from the target is 40 dB SPL, then the target strength is −40 dB. Note that this also takes into account the spherical spreading loss from the target, so it is important to use the same distance units across the sonar equation otherwise spherical spreading loss may be incorporated twice. Equations and nomograms exist for the calculation of target strength of ideal targets such as spheres and disks (*e.g.* Møhl 1988). For studies relating to the capabilities of bats this parameter must be measured for realistic targets. Typical target strengths for target moths with wingspans of 2 cm are of the order of −50 dB standardised to 1 m (Waters *et al.* 1995).

The final parameters in the sonar equation are the noise levels and the directivity index. Noise can come from many sources such as other bats, ultrasonically singing insects, wind noise in the ear of the bat as it flies forward, or spontaneous neurone discharge in the bat's cochlea. Some of this noise can be moderated by the directivity index: the model function of the bats hearing that in effect improves the overall signal-to-noise ratio. This directivity limits the direction over which the ear can receive noise. It can be defined as the dB reduction in the overall noise level which occurs as a result of the narrowing of the angle of view of the ear.

Both noise level and directivity index are difficult to estimate in non-idealised situations. As an example, a bat at 2 m from a moth having a target strength −50 dB using a call at 50 kHz (with an atmospheric attenuation component of 1.7 dB m^{-1}) and calling at 90 dB peSPL at 1 m is likely to receive an echo of:

$$90 - \left(4 \times 1.7 + \left(40 \times \log_{10}\frac{2}{1}\right)\right) - 50 = 21.2\,\text{dB SPL}$$

All values in the above equation have been normalised to a distance of 1 m.

The value of 21.2 dB SPL excludes any noise terms. At 4 m the value would be 2.3 dB SPL. If we assume the threshold of hearing is at 0 dB SPL (as it is for humans), then it can be seen that even with very high source levels, the range of echolocation is very limited.

The high output levels present a serious problem for bats since they may forward mask the very weak returning echo which must be detected a few milliseconds later. To overcome this, bats contract their middle ear muscles during echolocation to attenuate the outgoing signals and reduce the effect of forward masking.

6.3.2 Echolocation Call Structures

Bat echolocation calls can be divided into two types: frequency modulated (FM) and constant frequency (CF). FM calls are used by the majority of bats. CF calls are used by the Horseshoe Bat (*Rhinolophidae*), Leaf-nosed Bat (*Hipposideridae*) and the Moustached Bat (*Mormoopidae*).

FM signals are typically between several and several tens of milliseconds in duration, and sweep from a high frequency to a low frequency. The frequency and temporal structure of echolocation calls is intrinsically linked to the bat's habitat, preferred prey, and flight behaviour (Fenton 1995). A key influence is the effect of pulse-echo overlap (Schnitzler *et al.* 2003). Bats produce very loud outgoing pulses and are simultaneously listening for weak returning echoes. In addition they switch-off their hearing during their echolocation call. Hence they are insensitive to any echoes which return while they are producing their outgoing pulse. For a pulse of 5 ms duration, this means that any targets within 0.86 m are undetectable since the total path length for a 5-ms pulse travelling at 344 m s^{-1} is 1.72 m, which gives a target distance of 0.86 m.

Bats which fly high tend to use long duration, low frequency calls since they are in an open space with few pulse-echo overlap problems. They require the low atmospheric attenuation associated with low frequencies to probe long distances. Conversely, bats which feed close to vegetation tend to use very short duration signals, which contain more high frequencies to avoid pulse-echo overlap, and to resolve small features of their environment where attenuation is less of an issue. Bats will also alter their call structures as they home-in on targets to minimise the effects of pulse-echo overlap (Schnitzler and Kalko 2001).

When searching for prey, bats will typically produce *search phase* calls at a production rate of about 10 Hz, which are linked to their wingbeat cycle. Large bats with slower wingbeats produce fewer calls per second, while smaller bats tend to produce more. When the bat detects a potential prey item, it enters *approach phase*, where the repetition rate increases and the calls reduce in duration. Finally, just prior to the moment of prey capture, the bat enters the *terminal phase* where the call bandwidth reduces, the calls reduce below 1 ms in duration, and are produced at rates of up to 150 Hz.

Bats which use CF exploit a different technique to avoid the issue of pulse-echo overlap. They use a long duration (typically 50 ms) high frequency call with a rising FM sweep at the beginning, and a falling FM sweep at the end. Their cochlea has a highly sensitive region at the bats echolocation call frequency. However, outside of this range are areas of very low sensitivity. As the bat flies forward, it lowers the frequency of its call to one that falls into the area of low sensitivity, making it effectively deaf to the outgoing call. Any returning echo, however, is Doppler-shifted to a higher frequency as the bat is flying into the returning echo wave front. The bat manipulates the frequency of the outgoing signal such that the returning echo is placed within the area of high acoustic sensitivity, thus rendering the echo audible (Schnitzler 1973). This negates the problem of pulse-echo overlap, and has the additional benefit that the long duration returning echoes encode both amplitude and frequency modulations from the movements of the prey's wings, thus giving information on the prey type.

6.3.3 Signal Processing Capabilities

At its simplest the detection system of the bat can be thought of as a simple pulse-echo system. However, timing accuracy is crucial for accurate target location. An error of 1 ms translates to a target distance error of 17 cm. Since the typical time-course of a neurone action potential is of the order of 1 ms, any finer time resolution appears impossible. However, some experimental evidence suggests that bats may be able to resolve time to the order of a few hundred nanoseconds, equivalent to sub-millimetre range estimations. Temporal resolution of this precision, coupled with the ability to identify signals in noise, suggest a form of cross-correlation receiver model, whereby the returning echo is cross-correlated with a template of the original outgoing signal.

An FM signal produces a much sharper peak in the resulting plot than a CF signal, suggesting a reason why FM signals are preferred. However, a cross-correlation receiver requires the bats to be sensitive to phase at very high frequencies. Phase sensitivity is lost in humans at frequencies about 1.5 kHz, and there is no currently known mechanism by which phase at such high frequencies can be encoded.

One model which has been proposed to deal with these inconsistencies is the Spectrogram Correlation and Transformation receiver (SCAT) model (Saillant *et al.* 1993). This model uses a series of modular blocks to process incoming echoes. The first of these processing units is the *cochlear block*, where the incoming echo signal is bandpass filtered into 81 separate 3-kHz bands. The outputs from this filter-bank are then low-pass filtered to recover the envelope. This is followed by a *temporal block* in which the envelopes from each filter are passed into a series of parallel delay-lines. Neurones, tuned to specific delays in increments of 1 μs, scan for coincidences across the delay lines. The final block, the *transformation block*, is used to reconstruct target *shape* from multiple overlapping echoes derived from surface features of the target.

Recent experiments have shown that bats show very high levels of target resolution, to the extent of perceiving differences in the impulse response of complex targets (Grunwald *et al.* 2004). The impulse response of a target is the echo received when ensonifying the target with an acoustic impulse of infinitely short duration and infinite amplitude. Such a signal also has infinite bandwidth. In practise, a very short duration *click* stimulus provides a good approximation. The impulse response contains significant information on the target structure. Such experiments suggest that the bat may be able to deconvolve the impulse response of the target from its echo, though the mechanism by which this is done remains unknown.

6.3.4 Applicability of Bat Echolocation to Sonar System Design

Most systems that use ultrasound for navigation or orientation are based on simple pulse-echo designs. They benefit from understanding the constraints that bats experience in the physics of signal design and its transmission through the environment. As an example, higher frequencies provide more directional sound fields; coupled with higher target strengths from small objects, but with the costs of relatively greater signal attenuation and reduced range.

Engineered sonar systems also suffer from problems of pulse-echo overlap when close to the target, as do bats. One immediate area which may be of benefit in the design of bat-inspired assistive technology is binaural processing. Bats have one transmitter (their mouth) and two receivers (their ears). They are able to locate the position of an object in space by the echo delay, giving range; and by the relative intensity, phase and timing differences between the ears, giving the object vector.

By contrast, to sample multiple angles, sonar devices often use multiple transmitter and receiver pairs, each with a narrow angle of view. The decoding of multiple target vectors from a binaural receiver pair is obviously computationally complex, but is potentially of value in providing a more complex spatial map of the environment from a smaller number of transducers.

The main difference between sonar assistive devices and bat echolocation is the quantity of information extracted from the returning echo. Bats use the full range of fine resolution target information available to them in terms of spatial location of multiple targets and even target type and texture. Two issues currently prevent the incorporation of this information in an assistive technology device. First, many of the apparent pseudo-analogue, signal processing techniques used by bats to extract this information are unknown. It is of course possible to use conventional digital signal processing techniques to extract such information, such as cross-correlation and Fourier analysis, but this is computationally intensive and cannot currently be realised in real-time on a small portable device. The second, and potentially more limiting problem, is how to relay the wealth of information to the user in a meaningful way that preserves the full information spectrum, while not otherwise limiting the user's senses and ability to perceive their environment.

6.4 Design and Construction Issues

The design process embraced a range of key issues. There were based upon the specification of a set of outline requirements.

6.4.1 Outline Requirement Specification

The primary requirements for the aid were identified as:

- To exploit the confidence in the guide cane that users typically share. A further benefit of a system based upon a guide cane is that it clearly is able to provide a basic level of functionality in the event of failure of the technology for any reason.
- To design a sensory system that does not monopolize the hearing abilities of its user. Dependence upon acute hearing would limit the range of users who could benefit. For users who have a hearing capability it is important to allow this to be completely available to detect traffic and other hazards, and to simply allow conversation with others.
- To provide prioritized information on potential hazards and surrounding features to users within their 'navigation space', to enable them to move more confidently and effectively. A major benefit would be to extend the 'effective

distance' of the object detection capability relative to a standard cane, and to extend the 'space' to include potential hazards above and ahead.

- To exploit cognitive mapping to gain a capability that develops quickly with practice and also becomes subliminal to its user.
- To offer maximum convenience at a practical level: in terms of long battery life; of compact storage usability; and of flexibility of cane length and individual preference for cane tips.
- To provide a high quality design that is functional, pleasing and elegant to mature users, and appealing, stylish and desirable to younger users.

The primary performance requirements were translated into the twin core subsystems: the spatial sensing subsystem; and the user interface subsystem. These are integrated and controlled by a conventional embedded microprocessor. The cane should provide the functionality of a conventional guide cane and provide an elegant housing for power supply, processor, electronics, user-interface and sensor parts.

6.4.2 Ultrasonic Spatial Sensor Subsystem

A range of spatial sensing techniques were assessed in terms of range, coverage, and the availability and cost of sensor devices. The inspiration of the bat from nature shows ultrasound to be a powerful sensing capability that with appropriate processing can deliver high resolution spatial information to form the basis of assistive technology. The design process began with simple proof of principle trials. These explored the variety of the parameters of the ultrasound transducers, tactile feedback devices and their deployment.

6.4.3 Trial Prototype Spatial Sensor Arrangement

The resulting design employed a forward-looking ultrasound transducer, two transducers viewing respectively to the forward-left and forward-right, and importantly a further transducer looking upwards ahead of the user, to detect hazards that would otherwise threaten the head and neck area. The resulting electronic guide cane thus emulates the example of the bat through multiple transducers that detect targets from the 'navigation space' that the user may enter. Signal processing algorithms were needed to operate upon the ultrasonic data and perform a risk assessment whose results of potential hazards, in terms of criticality and distance, are available for delivery to the user.

In the early designs the transducers were located near the tip of the cane, except for the upward viewing transducer located on the handle. Figure 6.2 shows the schematic arrangement of the four transducers and their field of interrogation. A simple laboratory prototype based upon the above concept was designed at the University of Leeds. A short programme of feasibility study trials confirmed the potential of the concept.

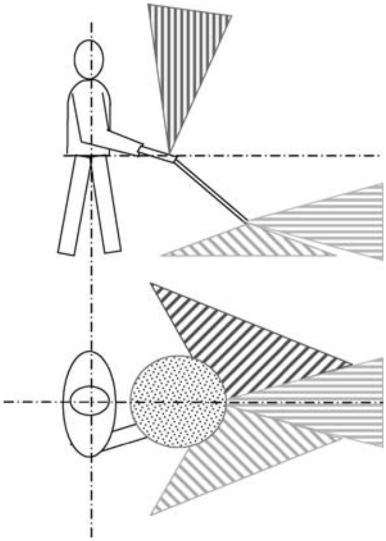


Figure 6.2. Schematic side and overhead views of transducer interrogation regions

Further development towards a concept evaluation prototype designed for use in a mass trial was next carried out in conjunction with Cambridge Consultants. This is shown in Figure 6.3. The left hand image shows the top of the handle and its upward viewing transducer nearest the shaft. It also shows in the background the forward and left field transducers in the cluster near the tip of the cane; the right field transducer is hidden.

Design of the underlying electronic subsystems is demanding for all portable devices where battery life is a key issue. Details are outside the scope of this paper, but the design addresses the efficient generation of the high voltage pulse required to excite each ultrasonic transducer in transmit mode. To optimize component utilization the same transducers are used for both transmit and receive modes.

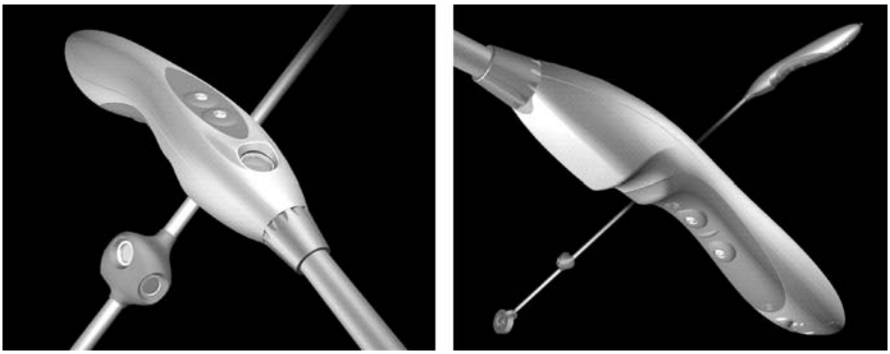


Figure 6.3. General views of the prototype trial cane, with detail of ultrasound transducers and tactors

To provide reliable detection of objects the transducers must interrogate the user space frequently and for efficiency share the pulse generation circuitry through a multiplexing system driven by an embedded processor. Received echoes from each transducer are also captured, encoded and passed to the processor for interpretation and integration with other transducer data to determine the hazard risk assessment.

6.4.4 Tactile User Interface Subsystem

The user interface needed to relay assessed hazard information is clearly a critical element. A number of inventions offer information to visually impaired users *via* audible signals, typically necessitating the wearing of headphones. Although this may be the most direct way to harness ultrasound echo information it is unacceptable as it denies this valuable sense to those having little or no vision. As noted above a core requirement is to retain unimpeded use of any hearing ability that a user may possess. This is obviously important in detecting surrounding moving hazards such as motor vehicles and people. Equally obviously, people who have a hearing impairment will be unable to use any aid that relies upon this to deliver guidance information. A multi-point tactile user interface has been designed to address this need. This has two key aspects: first, the design of the tactile transducers and their simple ergonomic matching to the human body; second, the cognitive mapping in the human brain that relays and interprets the tactile data to generate high level spatial awareness.

The use of a long cane in a sense allows its user to explore the world through an extended sense of 'touch'—exploiting the human capability noted above. This extended concept of 'touch' was harnessed in the electronic cane to present processed information from the ultrasound transducers. The human hand is one of the richest parts of the body in terms of tactile sensitivity. Thus a tactile transducer, or *tactor*, able to present sensory information to the fingers and palm surfaces of the hand, offers a highly efficient route to the sensory perception parts of the brain. The left hand image of Figure 6.3 shows the tactors on the top surface of the electronic cane handle. These are linked *via* the processor to the upward and forward ultrasound transducers. The right hand image shows the tactors on the bottom surface that are linked similarly to the left and right transducers.

Figure 6.4 illustrates the corresponding thumb and fingers positions and the support provided for right and left handedness. Handedness influences the ease with which tactile sensations are associated with a particular direction. The right and left tactors are simply inverted to exploit this linkage.

The tactors inform the user of an object in the field of view of the corresponding ultrasound transducer through a vibration whose frequency increases as the user approaches. A non-linear relationship is found to be optimal. The use of a frequency mapping is preferred as this is much less prone to fatigue and loss of sensitivity. The vibration is of course intermittent and only occurs when a hazard is in view of a transducer.

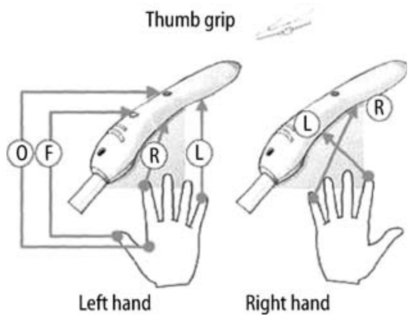


Figure 6.4. Prototype handedness arrangement

6.4.5 Cognitive Mapping

From the inspiration for the electronic cane, it is apparent that the bat builds a 'spatial map' of its surroundings using its ultrasound-based sensory capability. Human beings who have the sense of sight also build a 'map' of their surroundings. Physiologists have demonstrated that visually impaired people can also have an apparently identical capability, and this can also be present to some extent in those who have never had sight. In essence, human beings can typically visualise a scene in their mind. Thus information presented to a user, that includes appropriate range and direction information, may be implicitly used to form such a spatial map.

Various brain structures contribute to this ability. The midbrain structure, the *superior colliculus*, has the prime function to collate novel sensory input and organize the most appropriate motor response. It utilizes spatially aligned visual and auditory maps and the map of the body surface is superimposed over the same neural area (Stein and Meredith 1993). The electronic cane relies upon the sensory alignment in this brain structure by stimulating the near-space of the body map, in a spatially discrete manner. This is intended to give rise to an interpretation in the central nervous system of external spatial awareness.

The *somatosensory cortex* provides a further contributory proprioceptive knowledge of the position and attitude of our limbs. With practice a user will be aware of the orientation of a hand-held cane, and of the relative position in space of its extremity. Since the electronic cane has transducers to extend both range and dimension the new extents are similarly made available to allow the user to interrogate the space more precisely and complete the perceived spatial map. Clearly memory of known areas and their features will also provide further reinforcing information. The various factors that aid the cognitive process from the tactor stimuli are illustrated in Figure 6.5.

These cognitive processes are exploited in the electronic guide cane through the tactor devices to stimulate and harness the human perception system, and engender in the user's brain a spatial map of their surroundings, offering increased mobility and independence.

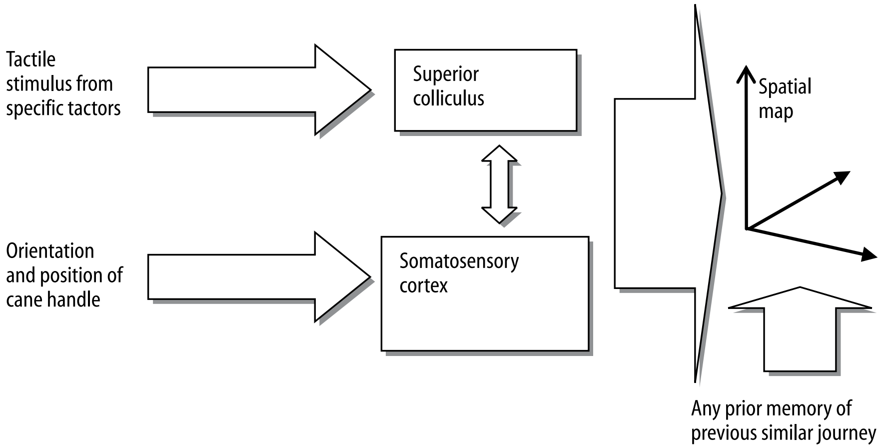


Figure 6.5. Cognitive pathways from tactor stimuli and handle position to derived spatial map

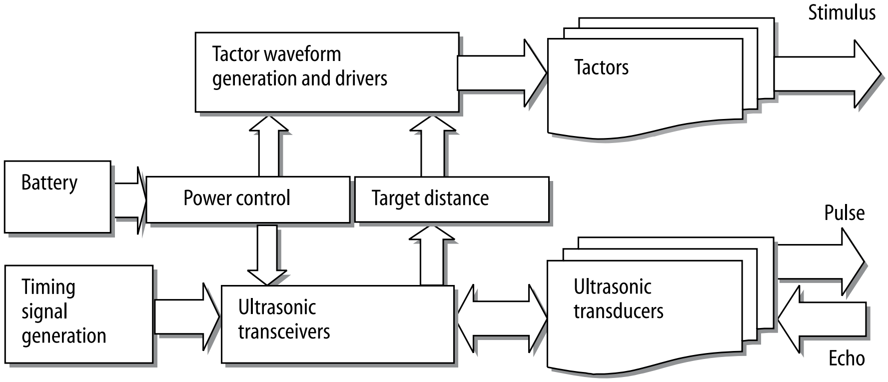


Figure 6.6. Major processing and control operations

6.4.6 Embedded Processing Control Requirements

The integration of the various signals and tactile transducers is achieved through a small embedded microprocessor. The processor also deals with the duty cycle control of key sensors and actuators to maximise battery life. Major processing operations are illustrated in Figure 6.6.

6.5 Concept Phase and Engineering Prototype Phase Trials

Early development work aimed to produce a concept testing unit built with no consideration for portability, power supply considerations, or optimized electronic design. A configurable processor with a connectable link to a desktop computer allowed the embedded software to be changed easily in order to investigate a variety of designs and algorithms.

A user review was incorporated at this concept phase to determine the viability of the basic concept. Results were very encouraging; approximately 10 concept trial users were found to gain a basic and intuitive competence within 30 min, and were able to demonstrate clear improvement in mobility in a controlled indoor environment featuring simple obstacles. Video recordings taken by agreement with the trial volunteers provided an opportunity for the self-learning and exploration sessions to be evaluated later.

The detailed engineering and trial prototype, illustrated above in Figure 6.3, was developed and manufactured in quantity to permit simultaneous trials by volunteer users to evaluate the electronic guide cane. The trial took place over a two-month period in which each volunteer used the cane in their everyday life. Trials were carried out over a six-month staged programme with volunteers in the UK, Germany and the USA. Volunteers were asked to keep detailed records of each use on their medium of choice. Records were used to assess the type of usage and the level of confidence. Statistics for the tests are given in Table 6.1.

Table 6.2 provides a summary of the results and typical user comments. As summarised in the table, the positive indications at the concept stage were confirmed and consolidated through the extensive experience of the trial period.

A small number of specific technical improvements were commonly indicated by the trial volunteers. At the trial prototype design stage it was expected that users would prefer a foldable form of the electronic cane for increased convenience. This

Table 6.1. Trial format, subjects and reviewers

Trial format and period
1 month using (2) forward and upward tactors
1 month also using side-view (4) tactors
Volunteer trial subjects
27 Visually impaired volunteers (54 person-months)
U.K., Canada, USA, Germany
Other reviewers and advisors
22 Orientation and mobility specialists
10 Market influencers

Table 6.2. Summary results and comments

73% – felt confident using within 1 week
53% – felt £360–£720 appropriate price range
Typical comments
“Keeps me constantly updated in a crowded area”
“Let me explore surroundings to a much wider degree”
“Positive guidance ... gave confidence ...”
“Knew when detecting – it was true and precise”
“Tells me things I wanted to know”

obvious feature was omitted in the trial form to reduce the design challenge and cost, and in order to concentrate attention upon novel aspects.

As expected practically all the trial volunteers pressed for the device to be foldable. A second improvement implied was the provision of short and long-range control, accompanied by changes to the forward transducer beam shape to more closely track the user's expected corridor of motion.

6.6 Case Study in Commercialisation

Results from the large scale trial summarized above were used to inform the optimization of the design in several areas: the incorporation of a foldable cane; the addition of a short- and long-range (to allow use in confined and open areas respectively) and accompanying changes to the transducer beams.

The incorporation of a folding capability was relatively simple in one respect: all the active components were re-sited to the handle assembly. This enabled the use of a standard multi-section graphite cane. It also necessitated the re-design of the transducers to suit their new position. Following the trial and its indicated changes, the needs and features of Table 6.3 were identified as targets for a product design.

The first 'Ultracane' product has been designed to include these features. For simplicity it includes twin ultrasonic transducers providing major forward and overhead directional information. The included angle, between the handle and cane shaft, allows the forward view transducer to be angled sideways to 'shoreline' a wall or other adjacent feature.

Two forward view distance ranges are provided, a long-range for use in open spaces and a short range for use in more crowded areas such as busy streets. The range can be changed at anytime. The selector is combined with the on-off control in a *push-pull* fingertip slider control on the underside of the handle. The position of the control provides a tactile confirmation of the current selected mode to the user. Figure 6.7 illustrates the resulting handle design and its features.

Table 6.3. Major identified need and proposed features

Users needs	Product feature
Enhanced mobility and decision making	Correct level of tactile, intelligent information
Ease of use	Intuitive, ergonomic, lightweight, minimise repetitive strain
Customisable	User choice of length and tip
Rugged	Hard shell, flexible components, graphite cane
Enhanced safety	Overhead sensor, 100% preview
Ease of learning	Mind mapping, builds on existing skills
Attractive and up-to-date	Desirable design with advanced technology
Affordability	Low purchase and cost of ownership *vs* competition

Figure 6.7. Ultra Cane® transducers and tactors

The left-hand image shows the handle from above with the handle-mounted overhead and forward transducers (and lobes for smaller sideways transducers not included here). The corresponding tactors are also visible.

Further details of the resulting commercialised product can be seen at www.ultracane.com.

6.7 Chapter Summary

The use of the long cane by visually impaired people as an obstacle detector is long standing. More recently the basic cane design has been equipped with laser or ultrasound transmitters and sensors and an interpretive human interface to improve its effectiveness, the objective being to allow safe travel by a visually impaired person.

This chapter reported a full and complete case study of the steps involved in developing an advanced technology obstacle avoidance cane that successfully used bat echolocation signal processing techniques and ultrasonic technology. The final cane design is now marketed worldwide as the UltraCane™.

The chapter began by reviewing the basic technological principles for ultrasonic waves and advanced signal processing methods. Thus, the scientific principles of the propagation, reflection and the collection of ultrasonic waves were presented. An extended presentation of bat sonar and the associated value of bat echolocation for obstacle detection followed.

The inspiration behind the first Batcane prototypes was to combine the use of ultrasonic technology with bat echolocation principles to obtain an efficiency gain in the ability of cane technology to detect obstacles. The chapter presented a detailed discussion of all the design and construction issues involved in creating and testing the first engineering prototypes. The use of vibrating buttons or *tactors* to indicate the location of an obstacle was an important feature of the cane user's interface.

The prospect of producing the cane as a commercial product soon began to emerge from the phase of developing and testing a satisfactory prototype. Consequently, the final part of the chapter examines the issues involved in bringing the prototype to eventual commercialisation. These included determining which features of the Batcane prototype should survive into the commercial UltraCane product.

Acknowledgement. The authors of this chapter have enjoyed participating in the design of the *UltraCane*, by Sound Foresight Ltd. See www.UltraCane.com for further information on its design and features. The authors also acknowledge the kind assistance provided by many organizations supporting visually impaired people in the UK, Germany, Canada and the USA and by the many volunteers who have tested prototype designs and enthusiastically provided encouragement and objective feedback. We also acknowledge the major product design contributions of Cambridge Consultants, Minima Design and Qinetiq at various stages.

Questions

Q.1 *On detection angles*

An experimental trial has determined that it is important to be able to detect a gap of 1 m between two objects (equivalent to detecting an open door) at a distance of 2 m. The transducers being used have a diameter of 32 mm. What is the minimum frequency of ultrasound that would give the angular resolution to detect the gap and not the door-frame?

As a first stage, use basic trigonometry to derive the angular width of the door at 2 m. Then rearrange Equation 6.4 to give the frequency that gives that angular beam pattern. A higher frequency would give a narrower beam. What is the problem associated with using higher frequencies?

Q.2 *On detection distances*

The user of a sonar device needs to detect a wall at 90° at a distance of 6 m. The transducers use a frequency of 40 kHz with an output level of 110 dB at 0.1 m and are sensitive enough to receive an echo at 32 dB SPL. Will the user be able to detect the wall?

First, find out the excess atmospheric attenuation at 40 kHz from Figure 6.1. Next, calculate the incident sound pressure at the wall using spherical spreading from Equation 6.6 and excess attenuation. Using the target strength of a planar target of −6 dB at 1 m, and spherical spreading and atmospheric attenuation on the return path, calculate the final echo strength received. Is it above the 32 dB SPL threshold of the receiver?

Projects

P.1 *On maximum and minimum detection ranges*

A useful exercise is to understand how parameters such as frequency, source level and the target strength of an object can affect detection distance. The best way of modelling these effects is to construct a simple spreadsheet which calculates the incident sound level at a target and the intensity of the reflected echo. To start, construct a column with target distances in increments of 0.1 m up to a maximum of 10 m. You will also need three cells to use to input the parameters of source level, atmospheric attenuation and target strength.

Next, use the distance values to calculate the incident sound intensity at each of the target distances and place these in column 2 next to the distances. To do this, you will need to consider the effect of spherical spreading and atmospheric attenuation. The effect of spherical spreading can be derived from Equation 6.6, using the source level you have input which should be standardised to a distance of 1 m. The atmospheric attenuation term can be added in as a simple product of the distance and the attenuation factor for the frequency you wish to consider. This can be read from Figure 6.1.

Next, use the incident sound levels you have calculated in column 2 as the source level for the returning echo. Subtract the target strength, and then use the same principles of spherical spreading and atmospheric attenuation as you did to calculate column 3. You should now have three columns. Column 1 will give you the distance to the target, column 2 will give you the incident sound pressure at the target for each target distance, and column 3 will give you the returned echo intensity for a target at each distance. You can now explore how source level, frequency (and attenuation) and target strength affect target detection distance by changing these parameters. If you have set up your spreadsheet correctly, changing these parameters will change all the received and echo intensities. You may have to decide on an arbitrary cut-off in sensitivity for your receiver device, such at 20 dB SPL. Hence, to determine the maximum range that a signal can be detected, look down column 3 until you reach the received echo intensity of 20 dB SPL, and then read across the distance. As a check, for an emitted sound level of 90 dB SPL at 1 m, a target strength of −20 dB, and attenuation of 1.5 dB m^{-1}, this should be around 6.1 m. For some starting parameters, the typical maximum output of a piezoelectric transducer is 110 dB SPL, with a maximum receiving sensitivity of 40 dB SPL at 40 kHz, where attenuation is 1.4 dB m^{-1}. The typical target strength of an object such as a 10 cm diameter street signpost is approximately −30 dB. For planar targets such as walls, assume the target strength is −6 dB at 1 m, since the major effect is the spherical spreading loss back from the target to the receiver.

P.2 *Building a bat detector*

The ultrasonic calls of bats are normally so far above the range of human hearing that we never detect them. However, several bat detector designs are available which convert the inaudible echolocation calls of bats into audible outputs. Most of these detectors work on a heterodyning process whereby mixing two signals of frequency f_1 and f_2 produces two outputs, $f_1 + f_2$ and $f_1 - f_2$. If the bat call is f_1, at around 40 kHz, then mixing it with another signal of 38 kHz produces an output of 40 + 38 = 78 kHz, which is also inaudible, and 40 − 38 = 2 kHz which is audible and can be heard *via* a loudspeaker. A 42-kHz signal mixed with the bat call produces a frequency of −2 kHz, which is 2 kHz with a 180° phase shift but acoustically identical to a 2-kHz signal. The mixing signal is generated by an oscillator whose frequency can be selected using a simple rotary control. This allows the user to scan up and down the frequency range used by bats to identify the maximum and minimum frequency present, as well as different tonal signals depending on the duration of the signal within the heterodyne window. There are numerous electronic circuit designs and kits for bat detectors on the Internet. For example a simple kit is available from http://www.magenta2000.co.uk. A further example uses a piezoelectric transducer coupled to a simple heterodyne circuit is available from http://www.alanaecology.com. Circuit designs are also available for frequency division detectors, where the received frequency is divided by a factor (usually a power of 2) and output to a tape-recorder or headphones. Frequency division detectors tend to be less sensitive overall due to their broadband sensitivity to noise, and are less useful for listening in the field, but are useful for analysis if the signals are recorded to tape and analysed later.

Listening to bats using their echolocation calls while watching them catch prey and detect and avoid obstacles gives great insight into the potential for sonar orientation and navigation.

References

Fenton, M.B., 1995, Natural history and biosonar signals, In: *Hearing by Bats* (ed. A.N. Popper and R.R. Fay), pp37–86, Springer-Verlag, New York

Griffin, D., 1958, *Listening in the Dark*, Yale University Press

Grunwalf, J., Schörnich, S., and Wiegrebe, L., 2004, Classification of natural textures in echolocation, *Proceedings of the National Academy of Sciences*, Vol. 101, pp. 5670–5674

Møhl, B., 1988, Target detection by echolocating bats, In: *Animal Sonar: Processes and Performance*, (ed. P.E. Nachitgall and P.W.B. Moore), pp. 435–450, Plenum Press, New York

Morse, P.M., 1948, *Vibration and Sound* (2nd Ed.), McGraw-Hill, New York

Rossing, T.D., and Fletcher, N.H., 2004, *Principles of vibration and sound* (2nd Ed.), Springer-Verlag, New York

Saillant, P.A., Simmons, J.A., and Dear, S., 1993, A computational model of echo processing and acoustic imaging in frequency-modulated echolocating bats: The spectrogram correlation and transformation receiver, J. Acoustical Society of America, Vol. 94, pp. 2691–2712

Schnitzler, H.U., 1973, Control of Doppler shift compensation in the Greater Horseshoe Bat, *Rhinolophus ferrumequinum*, J. Comparative Physiology, Vol. 82, pp. 79–92

Schniztler, H.U., and Kalko, E.K.V., 2001, Echolocation by insect-eating bats, Bioscience, Vol. 51, No. 7, pp. 557–569

Schnitzler, H.U., Moss, C.F., and Denzinger, A., 2003, From spatial orientation to food acquisition in echolocating bats, Trends in Ecology and Evolution, Vol. 18, pp. 386–394

Stein, B.E., and Meredith, M.A., 1993, *The merging of the senses*, MIT Press, Cambridge, Mass. USA

Waters, D.A., and Jones, G., 1995, Echolocation call structure and intensity in five species of insectivorous bats, J. Experimental Biology, Vol. 198, pp. 475–489

Waters, D.A., Rydell, J., and Jones, G., 1995, Echolocation call design and limits on prey size: a case study using the aerial-hawking bat *Nyctalus leisleri*, Behavioural Ecology and Sociobiology, Vol. 17, pp. 312–328

Wells P.N.T., 1969, *Physical principles of ultrasonic diagnosis*, Academic Press

Resources

www.UltraCane.com – more information about the UltraCane and its development

7 Navigation AT: Context-aware Computing

Learning objectives

In order to allow visually impaired people the ability to travel freely through the environment, without being constrained to familiar routes and known destinations, this chapter illustrates the need to augment local information acquired from commonly used mobility aids with distant information acquired from context-aware technologies.

When you have read this chapter you should be able to:

- List the types of commonly used mobility aids and explain the limitations of their use.
- Define the concept of cognitive mapping and describe the strategies that are used to acquire spatial information, and the factors influencing how this is acquired.
- Outline technologies that can be used to sense the context of the user, and discuss the technological and human issues that arise.
- Carry out test procedures involving end-users in order to investigate specific human issues affecting the design of context-aware systems.
- Discuss the technological challenges concerned with improving the accuracy and reliability of future positioning technologies.

7.1 Defining the Orientation/Navigation Problem

The ability to orientate and navigate is an important skill that is used to experience and interact with the environment, to make social contact with other people, to undertake daily activities, and, ultimately, to maintain *independent* mobility. In this section, we will look at:

- What orientation, mobility and navigation mean.
- What types of mobility aids exist and how they are used.
- The limitations of traditional mobility aids.

7.1.1 Orientation, Mobility and Navigation

'Orientation' refers to a person's awareness of his/her position in space. For instance, a visually impaired person may be aware of their current position due to the sound of a verbal travel announcement when passing a train station. Orientation has been defined as the 'process of utilising the remaining senses in establishing one's position and relationship to all other significant objects in one's environment' (Hill and Ponder 1976). Orientation is therefore dependent upon the solicitation and interpretation of sensory information, which may be visual, auditory, kinaesthetic (the relative positions and movements of different parts of the body), tactile, thermal, and/or olfactory. Successful interpretation of sensory cues is dependent upon known and predictable environments (LaGrow and Weessies 1994).

Orientation and Mobility (O and M) specialists teach visually impaired travellers to recognise and anticipate the regularities of the environment. However, exceptions to those regularities, such as the smell emanating from a bakery, become more informative than the regularities themselves, and become landmarks that a traveller can use to pinpoint their location in space (Hill and Ponder 1976).

'Mobility' refers to the ability of a person to move safely and efficiently from one *point* to another within the physical environment. It has been defined as the 'capacity, the readiness, and the facility to move' (Hill and Ponder 1976). This will involve sensing and negotiating obstacles and hazards, establishing and maintaining a desired course, and recovery from veers and other unintended or unexpected changes in direction. For instance, a visually impaired person may need to negotiate cars parked on the pavement, potholes in the road, temporary road signs, and excavation work.

'Navigation' refers to the purposeful process involved in travelling from one *place* to another, using mobility skills, and orientation in the environment in relation to a desired course. Navigation therefore involves the traveller updating their orientation and position, which can be achieved using three methods classified on the basis of object motion or kinematic order (Loomis *et al.* 2001):

1. 'Position-based navigation' depends on external signals within the environment, such as landmarks, indicating the traveller's position and orientation.
2. 'Velocity-based navigation' (normally referred to as 'dead reckoning') relies on external and internal signals indicating to the traveller their present position by projecting course and speed from a known past location, and predicting a future position by projecting course and speed from a known present position.
3. 'Acceleration-based navigation' (normally termed 'inertial navigation') involves both the traveller's linear and rotary accelerations to acquire information on displacement and heading change from the origin.

With respect to these three methods, visually impaired people are at a huge disadvantage in unfamiliar routes, as they 'lack much of the information needed for planning detours around obstacles and hazards, and have little information about distant landmarks, heading and self-velocity' (Loomis *et al.* 2001).

7.1.2 Traditional Mobility Aids

Traditional mobility aids have acted as an important and effective tool for helping visually impaired travellers detect objects in the local environment, negotiate narrow spaces, climb and descend stairways, enter and exit buildings, as well as many other mobility tasks. Three widely used traditional aids exist, and are described below.

Human guide. This is where a person with sight serves as a guide to a person who is visually impaired. The human guide is positioned slightly in front in order to ensure safety. At one time or another, most visually impaired travellers use a human guide, either as a primary aid or supplement to other aids.

Long cane. Over the years, various types of canes have been developed for specific user needs and preferences. The long cane (also called the prescription cane or typhlo cane); however, is the most effective and efficient (Farmer 1980). It enables visually impaired travellers to detect obstacles or drop-offs in the path approximately 1 m in front of them. Information regarding the walking surface or texture can also be transmitted, while providing suitable lower-body protection. In the most common technique, the cane is extended and swung back and forth across the body in rhythm with the user's steps (LaGrow and Weessies 1994).

Dog guide. Trained dogs are used as travel aids by less than 10% of non-visual travellers. The dog responds to commands given by the visually impaired person, such as right, left and forward (commands are only disobeyed to avoid danger, such as refusing to proceed forward due to a car parked on the pavement obstructing their path). The guide dog's job is not to find the way, but to guide the person around obstacles or stop in front of them. People who use guide dogs must therefore know where they are going and make decisions about the proper time to begin a street crossing.

The requirements and abilities of visually impaired people vary considerable, and so an O and M specialist would advise on a mobility aid (along with techniques and instructional strategies) that reflects the uniqueness of each person. The traveller's quality of vision, for instance, is a key factor in identifying which mobility aid is most appropriate.

7.1.3 Limitations of Traditional Aids

Long canes and guide dogs have been effective in helping visually impaired people with many mobility tasks, as described in the last section. However, Clark-Carter *et al.* (1986) state that at least one-third of people with visual impairment or blindness make no independent journeys outside their homes, and most of those who do venture outside independently often travel to *known* destinations along *familiar* routes, as exploration is considered stressful and can lead to disorientation. The inability of these aids to facilitate distant (or macro) navigation is considered the main reason for this (Petrie 1995). The visually impaired traveller, for instance, would remain unaware of a supermarket located at the other side of the street. Even for local (or micro) navigation these devices are limited as they only detect objects below waist height (see Chapters 5 and 6 for more discussion on obstacle avoidance technology and issues).

Very limited contextual detail of the environment can be acquired from traditional mobility aids. The focus of this chapter is therefore on assisting the visually impaired traveller for distant navigation by providing a greater spatial and contextual orientation beyond the immediate environment or what can be detected using a mobility aid. Visually impaired travellers will therefore not feel restricted to frequently travelled routes, as they will be supported in travelling to *unknown* destinations, along *unfamiliar* routes.

7.2 Cognitive Maps

Humans undertake many types of physical actions and activities in their daily lives, such as travelling to work, attending business meetings or social engagements, going on holiday, *etc*. The cognitive decisions or choices underpinning these spatial behaviours are based upon previously acquired spatial understandings of the world and perceived external cues or references (such as maps or street signs). 'Cognitive map' is a term which refers to 'an individual's knowledge of spatial and environmental relations, and the cognitive processes associated with the encoding and retrieval of the information from which it is composed' (Kitchin and Blades 2002).

Cognitive mapping research focuses upon how individuals acquire, learn, develop, think about and store data relating to the everyday geographic environment, such as locations, attributes, and landmark orientations to navigate (Downs and Stea 1997). The motivation behind this research is for *understanding* and *predicting* spatial behaviour by identifying the correlation between people's environmental representation with their behaviour in the environment.

Over the years many researchers have attempted to conceptualise cognitive mapping. Several complex models and theories have been proposed, some of which originate from geographical research, others from psychological theories, and more recent theories that incorporate both geographical and psychological principles. Haken and Portugali (1996), for instance, propounded the inter-representational network (IRN) theory, which emphasises the interdependence of internal (cognitive) representations and external (environmental) representations. For instance, when a person experiences a new environment there will be an interaction between internally stored representations derived from previous environments and the perception of external patterns in the new environment. These internal and external inputs create a cognitive map. IRN embodies principles from:

- Gibson's (1979) perceptual theory where it is argued that environmental features are encoded directly from perception without additional cognitive processing.
- Information processing theories (such as Golledge and Stimson 1987) which concern the flow of information between the individual and environment; the perceptual filtering of information; the factors that influence the interpretation

of, and decisions regarding, perceived information; and the revealed spatial behaviours.

- *Experiential realism* in that the patterns of cognitive processing are derived from the person's experience in the environment.

The remainder of this section will discuss the processes of, and the factors that influence, the acquisition of spatial information. The chapter then discusses computing technologies that support orientation and distant navigation by creating or augmenting a person's cognitive map. This leads onto the research area of context-aware computing, which concerns the integration of mobile technologies in order to transmit personalised (tailoring information to the user's needs and preferences) and localised services to the user when travelling through diverse environments. Methods for capturing key human design issues are then discussed in relation to human-computer interaction and cognitive mapping. The last section discusses expected and required improvements in location precision technology in the future.

7.2.1 Learning and Acquiring Spatial Information

According to research with sighted people, the strategies for learning spatial information can be considered from two different perspectives (comparable research with visually impaired people is in its infancy). First, *navigation*-based learning is where spatial information is collected and processed directly from the individual's interaction with the environment. Kitchin and Blades (2002) outline three main theories about how people learn an environment from spatial interaction:

1. *Landmark* theories, *e.g.* Golledge's anchor-point theory (Golledge 1978), are where environmental cues lay the foundation to which further information is added, such as the spatial relationship of landmarks in a path.
2. *Route* theories, *e.g.* Gärling *et al.* (1981), are the opposite of landmark theories in that path-based information lay the foundation to which spatial positions of landmarks along this path are added.
3. Theories concerning *ordered views/scenes, e.g.* Cornell and Hay (1984), suggest that way-finding can be dependent on memorising ordered views or scenes rather than learning landmarks and paths.

The second form of spatial learning is *resource*-based where spatial information is collected and processed without having to directly experience the environment. Resource-based learning can be acquired from atlases, television, newspapers/magazines, schooling, talking to others, and written and verbal directions. The process of acquiring this information, however, can be different for visually impaired people who use tactile maps, Braille newspapers, and embossed pictures to learn from resources that require sight. This type of learning is 'a useful supplement to direct experience, and is the only source of information about environments at scales that cannot be experienced directly, such as countries or continents' (Kitchin and Blades 2002).

7.2.2 Factors that Influence How Knowledge Is Acquired

Navigation-based and resource-based learning are influenced by various factors, all of which can be classified under two separate headings relating to *environmental* and *individual* variability. Environmental variability is addressed by Jonsson (2002) who describes how spatial information can be encoded differently depending on (i) the time of day, *e.g.* landmarks can appear differently at night, (ii) the type of season, *e.g.* snow in winter *vs* a summers day, (iii) the weather conditions, *e.g.* rainy day *vs* sunny day, and (iv) direction of travel, *e.g.* the appearance of landmarks change when travelling the same route forward and then back.

Individual variability is addressed by Kitchin and Blades (2002), who described how influencing factors may include gender, age, education, culture, emotion, beliefs, preferences, and abilities/disabilities. There is evidence, for instance, that elderly people have poorer spatial memory and spatial ability, *i.e.* the ability to process information about the relationships among objects in space and time. Gender differences in spatial ability have also been found. In small-scale tasks involving mental rotation and spatial perception, males perform better than females (Allen 1999). However, it is not known how important these abilities are in the development of cognitive maps.

The influence of disability on learning is a much-needed area for further research. Of the limited studies that have been undertaken, most researchers have focused on visual impairments, while others have carried out studies with wheelchair users, and people with neuropsychological and learning impairments. People without or with minimal vision, for instance, rely on sequential learning using tactile, prioceptive, and auditory senses to encode spatial information and construct spatial relationships (Bigelow 1996). There is limited research, however, into the acquisition of spatial information by people with varying degrees and forms of visual impairment. These types of issues are illustrated in Figure 7.1a–d.

So key questions relating to Figure 7.1 would include: how would people experiencing impairments similar to Figure 7.1a–d encode spatial information, which types of sensory receptors would be used to acquire different types of spatial information, and with respect to navigational aids, what assistance or information could be provided to enhance their spatial orientation or cognitive map? For instance, someone experiencing the advanced cataract condition shown in Figure 7.1c may be more dependent on encoding auditory information than Figure 7.1b,d, as objects are less distinguishable. Further, a navigational aid would need to provide assistance on textual features in the environment (such as street signs) for someone experiencing a loss of central vision, shown in Figure 7.1b, as reading text would be problematic.

Some have argued that by improving the design of built environment to be more accessible and memorable (Golledge and Stimson 1997) would facilitate the development of visually impaired peoples' cognitive maps. However, this will not tackle the problem of macro-navigation, as discussed earlier. Overall, more cognitive mapping research is required in order to reveal what spatial information should be given to visually impaired pedestrians, in what form, and at which particular locations (Kitchin and Jacobson 1997; Kitchen *et al.* 1997).

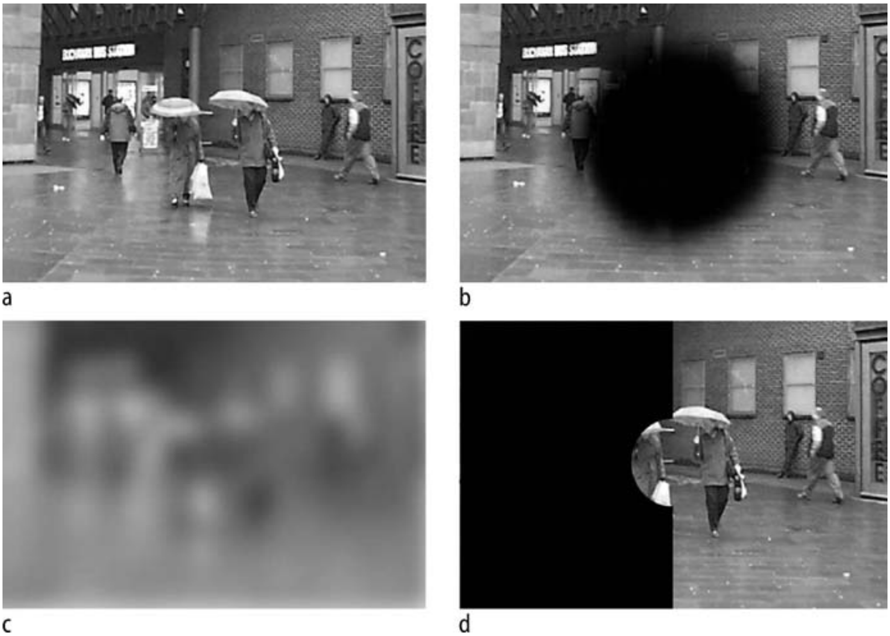


Figure 7.1a–d. Photographical representations of different visual impairments: **a** normal vision; **b** loss of central vision (this can be caused by macular degeneration); **c** possible effect of advanced cataract; **d** one half of the field of vision lost (may be due to stroke or head injury) (note: representing human vision pictorially is difficult, as binocular vision is three-dimensional and consists of focal and peripheral vision)

7.2.3 The Structure and Form of Cognitive Maps

Over the years there have been several theories proposed to account for (i) how cognitive maps are structured and composed, *i.e.* non-hierarchical, hierarchical, and schema theories, (ii) the form of, and mechanisms supporting, cognitive maps, such as images, dual coding, genetic coding, *etc.*, (iii) the process with which spatial knowledge is accessed and utilised, and (iv) how spatial knowledge is expressed (Kitchin and Blades 2002)

Jonsson (2002), for instance, differentiates between *active* and *passive* cognitive maps. Active maps contain spatial information that is always available and which can be described verbally, *e.g.* giving detailed directions to a disorientated tourist. In contrast, passive maps contain landmarks that are only recognised when the traveller sees them, *e.g.* revisiting landmarks after a long absence—returning to a former residence, holiday destination, *etc.*

In relation to (iv), there is evidence to suggest that people with visual impairments express their spatial knowledge differently to that of sighted people. Bradley and Dunlop (2003) found that visually impaired people provide richer contextual descriptions (when describing a route) including information not used by sighted participants, such as sensory and motion-based information. In a further study by Bradley and Dunlop (2004), significant differences were found between people with different visual impairments. When asked to walk to pre-determined outdoor and indoor landmarks, participants experiencing a central vision loss and total vision loss, for instance, asked more questions relating to side streets, steps, distance, and temporary obstacles, than participants with a peripheral vision loss. Such differences in the form and structure of spatial knowledge will need to be addressed if way-finding technologies (such as those illustrated in the next section) are to be usable by visually impaired people.

7.3 Overview of Existing Technologies

In the 1940s, the long cane was adopted as the primary mobility aid within the blind community (Farmer 1980). Since then, electronic travel aids (ETAs) have been developed in order to provide more detailed feedback regarding obstacle avoidance and navigation within the immediate environment (see Chapters 5 and 6 for more discussion on obstacle avoidance issues). However, in order to support independent mobility of visually impaired people to *unknown* destinations along *unfamiliar* routes, these devices will need to be supported with distant navigation technologies. This is the topic of the next section in which both indoor and outdoor technologies will be illustrated. The last section will describe output technologies for communicating information to the visually impaired traveller.

7.3.1 Technologies for Distant Navigation

Within the last decade, the development of ETAs has been guided more to the challenging topic of supporting distant navigation and independent mobility. One approach is to use a network of location identifiers that can be remotely sensed by the visually impaired traveller within either indoor or outdoor environments. Some systems use infrared transmitters installed throughout the environment to transmit digital speech about the location (see Chapter 10), while others use radio frequency (RF) beacons to wireless transfer digitally coded and compressed speech, *e.g.* Kemmerling and Schliepkorte (1998). The main drawback of this technology is the cost of installation. In addition, contextual information is limited (vital information regarding temporary obstacles is not included) and generic for all users. Individual requirements such as those described in Figure 7.1 would not be supported.

Global positioning system (GPS)-based navigation aids, combined with geographical information systems (GIS), have been used by many researchers to assist navigation for visually impaired people in outdoor environments (see Chapter 8). Here, signals picked up from satellites orbiting the earth's atmosphere are used by a mobile device to convert latitude and longitude coordinates to a geographical location using a digital map. Good examples of research-based GPS systems (other than Chapter 8) include: (i) the MOBIC Travel Aid which integrates differential correction from ground base stations (DGPS) for greater location precision, a compass worn on the body for heading direction, and mobile telecommunication facility (Strothotte *et al.* 1996), and (ii) a navigation system for the blind that uses a telephone connection to transmit the user's GPS coordinates to a central server, which sends back digitised speech to the visually impaired traveller (Makino *et al.* 1997). There are, however, privacy issues associated with the latter system, concerning the transmission of location information, that would need to be given further investigation.

While GPS capability offers great potential, its accuracy and reliability is still not good enough for safe and effective navigation by visually impaired people since DGPS is not available in many locations, and the signal from satellites can become blocked in narrow built up areas. In addition, most GPS-based systems have based the level of service solely around the GPS function, and have not considered other *contextual* information in the visually impaired person's environment that could facilitate and enhance independent mobility and navigation within unfamiliar environments (discussed further in Section 7.4). Helal *et al.* (2001), for instance, describe how many GPS-systems 'lack dynamic query capabilities and support for dynamically changing environments' and that 'context-awareness is not well supported'.

Unfortunately GPS is ineffective inside buildings and so most location-aware systems within indoor environments depend on relative positioning using various technologies. Active Badge, for instance, is a small wearable device which transmits a unique infrared signal every 10 s (Want *et al.* 1992). These signals are detected by one or more networked sensors, which are used to determine the location of the badge on the basis of information provided by these sensors. There is also evidence to suggest that location can be inferred using other methods by integrating information from accelerometers, magnetometers, temperature and light sensors (Golding and Lesh 1999) or using wireless networks (see Section 7.4).

7.3.2 User Interface Output Technologies

Most existing mobile GPS systems transmit information to the user by a visual display (or screen) only. However, those designed specifically for visually impaired users normally give information by speech audio. There are, however, notable problems using such an approach. Franklin (1995) highlights the difficulties of interpreting spatial relations from common speech, Pitt and Edwards (1996) indicate how speech interfaces are slow to use and more demanding on memory than vision or touch, and Strothotte *et al.* (1996) demonstrate how headphones used to transmit audio messages may mask/distort environmental sounds that visually impaired people use to avoid hazards. Other styles of communicating information include:

- *Non-speech output.* This is where audio signals, such as beeps, are used to direct the user to a landmark without the use of speech. Some researchers have studied the use of panning to indicate spatial relations, where contrasting tones indicate direction and the rapidity of sounds indicate distance (Holland and Morse 2001).
- *Haptic feedback.* Vibration alerts can be transmitted using tapping interfaces to indicate when the user needs to turn left or right. Refreshable Braille displays

could also be used, though this style is not preferred since the speed at which braille is read is often to slow for the rate at which objects are encountered in the environment. Also, only a very small proportion of the blind community actually read Braille.

Some research has looked into combining different output technologies. Ross and Blasch (2000), for instance, found that the most effective interface combined tactile cues using a tapping interface with improved speech output. More research and studies, however, are required to identify different preferences for communication styles.

7.4 Principles of Mobile Context-aware Computing

In the last section, examples of current technologies to support distant navigation for visually impaired people were described. Here, the principles of context-aware computing, where the notion of 'context-awareness' concerns the ability of a computing system to recognise a *user's context* and respond to it in a way that is useful to the user, are discussed. This ability is achieved by integrating information acquired from both multiple sensing technologies and other contextual resources, such as web-based servers and personal diaries/calendars. By integrating and then interpreting this information, the application becomes more aware of the user's environment and is able to provide more useful information and services in line with the user's task, location, and situation. Context-awareness could therefore support independent mobility of visually impaired people as it would augment their perception and understanding of the environment (Helal *et al.* 2001).

Adapting or personalising information to the individual user is also at the heart of context-aware computing. So, for instance, unique cognitive mapping strategies used for encoding spatial information could be accounted for by having more customised feedback. Using the example of people with different visual impairments (discussed in Section 7.2.1), it is vital that feedback is structured in such a way in order to allow people with different requirements to learn and experience the environment rather than become dependent on verbose and generic application feedback.

The capabilities of context-awareness become even more prominent when one considers the current trend towards ubiquitous or pervasive computing. This is an ideology originally propounded by Wesier (1991) who envisaged a world where computers are embedded in everyday objects allowing contextual information to be exchanged in an interconnected environmental infrastructure. As sensors become cheaper and smaller, this notion becomes a closer reality, allowing many new services to be available to the user. However, such advancements need to be investigated alongside the human and social implications of acquiring more information about the user, particularly privacy and security issues.

The remainder of this section will be discussed under five headings: (i) adding context to user computer interaction, (ii) acquiring useful contextual information, (iii) capabilities of context-awareness, (iv) application of context-aware principles, and (v) technological and usability issues.

7.4.1 Adding Context to User-computer Interaction

Most traditional desktop and mobile computers rely on explicit user actions either to provide task-based information or execute task-based services. This style of interaction is illustrated in Figure 7.2.

Computers that act only on explicit user input are context independent, since they are unable to adapt to the surrounding environment (as depicted in Figure 7.2). In other words, these computers are unaware of who or what is surrounding the user. Output is therefore determine by the commands given by the user, resulting in the user being caught up in a loop. This style of interaction is not natural for many modern computer settings and can be time-consuming, a demand on the user's attention, and likely to lead to user frustration caused by a mismatch between what the computer is capable of doing, and what the user would like it to do given their current situation or context. For instance, some users may desire a mobile phone that automatically changes its settings to silent when in a cinema, and gives access to that and nearby cinemas' programme guides rather than a standard web access to all cinemas in the country. Additionally, when considering human-human communication, traditional computers force users to interact and behave in a way that is unnatural to them. The content and nature of human dialogue, for instance, is verbalised in a way to suit the linguistic, social, task, physical, and cognitive context (Bunt 1997). In the case of mobile computing, these issues will be particularly important, where users move through complex and dynamic environments involving a myriad of interactions with other people and objects.

Context-aware computing is centred on the premise that, by adding context, interaction will become more natural and personalised. Further, traditional user-driven interaction will be minimised as some user actions could be inferred. This

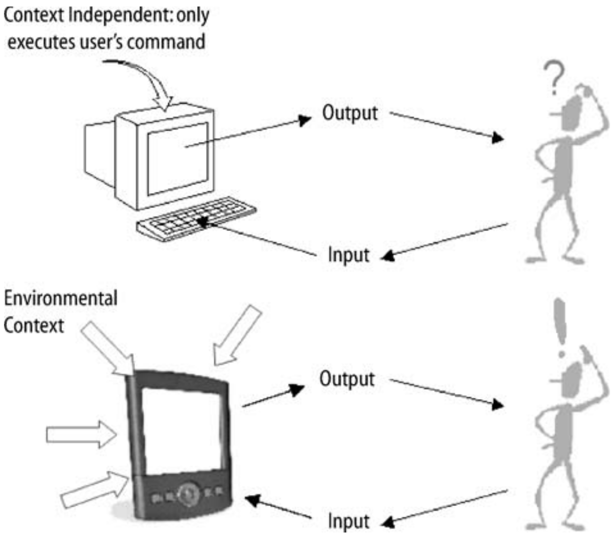


Figure 7.2. Interaction with traditional and context-aware user interfaces

last issue is particular relevant to visually impaired people, who would wish to focus their attention on hazard identification and environmental learning and experience, rather than frequent interaction with the device.

Context-aware computing therefore completely redefines the basic notions of interface and interaction (Moran and Dourish 2001), and 'by improving the computer's access to context, we increase the richness of communication' (Dey and Abowd 2000a). Examples of useful contextual information and the processes with which information could be acquired will be described in the next section.

7.4.2 Acquiring Useful Contextual Information

Before useful contextual information can be identified, it is important to firstly specify what is meant by the notion of *context*. This is a 'powerful, and longstanding, concept in human-computer interaction' (Moran and Dourish 2001), which provides an insight into how people behave, make decisions, select goals, and interact with other people and objects within their environmental context. Context is a complex and multidimensional term that has received many contrasting definitions and categorisations over the years. Typical dimensions of context are described in Table 7.1.

In other examples, Dey and Abowd (2000a) define context as 'any information that can be used to characterize the situation of an entity—an entity is a person, place, object that is considered relevant to the interaction between a user and

Table 7.1. Dimensions of context

Dimension	Definition
Physical	The environmental location consisting of surrounding/nearby physical objects (*e.g.* buildings, cars, trees, *etc.*). This also includes the presence, state and purpose of those objects, and the types of information they transmit through audio, visual, odour, texture, temperature, and movement (as well as under different weather conditions)
Social	The relationship with, and the density, flow, type, and behaviour of, surrounding people (*e.g.* sitting on a crowded train)
Task	The functional relationship of the user with other people and objects, and the benefits (*e.g.* resources available) or constraints (*e.g.* time pressure) this relationship places on the user achieving his/her goal
Temporal	The temporal context is embedded within everything, and is what gives a *current* situation meaning, based upon *past* situations/occurrences, expected *future* events, and the higher-level temporal context relating to the time of day, week, month, or season
Application	The capabilities and limitations of both the application (such as battery usage life, processor speed, memory capacity, sensors, input/output technologies, *etc.*) and the sources from which data is derived (such as the processing speed of a web-based server)
Cognitive	A user's cognitive processing abilities, short- and long-term memory abilities (containing past experiences), dislikes/preferences, opinions/beliefs, cultural interpretations, perceptual abilities (using five senses), cognitive mapping strategies for encoding spatial information, *etc.* In relation to the task, the user's cognitive context also consists of the high level goal along with the planned/expected structure, perceived timing, and composition of lower-level goals and physical actions

an application, including the user and application themselves'. So, here, a blind person's guide dog may be considered a relevant entity to the interaction between the user and the GPS-based navigation device, since the activities supported by the guide dog would not need to be supported by the device.

Some researchers have also proposed a multidisciplinary approach to context-aware design. Selker and Burleson (2000) describe how cognitive science needs to be integrated, while Bradley and Dunlop (2003) combine theories of context within linguistics, computer science and psychology. In the latter case, it is argued that adding contextual theories within (i) linguistics provide a useful insight into how contextual detail might be altered for different contexts and situations (Bunt 1997), (ii) psychology provides a invaluable insight into the cognitive processes that underpin a user's decisions and behaviour, such as cognitive mapping strategies (as discussed before), decisions regarding meaningful *vs* incidental aspects of the environment (Smith 1988).

Thus, an understanding of context will enable application designers to identify which dimensions are useful for both inferring a user's activity/intentions and providing more relevant context-specific feedback. Useful contextual information, all of which has been used by context researchers at some point, may include:

- *User's profile*. Preferences/dislikes, *e.g.* customised speech output settings.
- *Physical environment*. Location and nearby objects, *e.g.* in bus station and ticket machine nearby.
- *User's activity*. Orientation and speed, *e.g.* asleep or awake, on a bus or walking.
- *Temporal*. Time of day, week, or season, *e.g.* just finished work and a winter's day.
- *Environmental parameters*. Temperature, light, noise, and weather conditions, *e.g.* cold, dark, and icy.
- *Social environment*. People nearby, *e.g.* streets busy prior to a football match.
- *Resources*. Nearby and available, *e.g.* train and bus timetables.
- *Physiological*. Blood pressure, heart rate, and tone of voice, *e.g.* heart rate indicating level of anxiety/stress.

Contextual information can be acquired using various sensing technologies and contextual resources. Sensing technologies include GPS and Active Badge to identify user's location; ultrasonic transducers and Bluetooth to detect objects in the local environment; temperature sensors; physiological sensors to measure, for instance, heart rate to give an indication of anxiety levels; and camera technology and image processing for face recognition. Contextual resources, such as weather forecasts, can be acquired by connecting to web-based servers, while a user's diary and personal settings can be used to identify a user's planned social or business arrangements, holidays, activities, *etc.* While these sensing technologies can be used to better understand the user's context, the human implications of their design and use need to fully investigated, such as the intrusiveness of physiological sensors.

In order to make accurate inferences regarding a user's context, acquired contextual information will need to be effectively interpreted, consolidated and managed. This particular aspect of design is termed middleware design, which aims to provide toolkits and services to facilitate the discovery and delivery of context data in an efficient manner. Dey and Abowd (2000b), for instance, developed a component-based framework, called the Context Toolkit, which is centred on three main abstractions. First, 'context widgets' wrap around underlying sensors and deliver information to interested components/services, such as a change in GPS location delivered to an application that works out the user's distance from their destination. Second, 'context aggregators' store low-level logically related sensed data (such as a person or location) for relevant application entities, and, third, 'context interpreters' are responsible for abstracting low-level data to higher-level information. Some of the technological issues associated to this challenging area of research and development, will be discussed in Section 7.4.5.

7.4.3 Capabilities of Context-awareness

The capabilities of a context-aware system are dependent on both the technologies of the system itself, and the dissemination of, and level of communication with, other computers within a ubiquitous computing environment. For instance, a greater density and variety of computers embedded within everyday objects may result in (i) a greater exchange of contextual information, (ii) a greater awareness of the environment, and (iii) a more robust level of service for the user. Four categories of context-aware capabilities are described below (Pascoe 1998):

- *Contextual sensing*. Application senses environmental states and presents them to the user. For instance, a GPS receiver takes in a location, compares it to a digital map, and then informs the user of their location.
- *Contextual adaptation*. Application leverages contextual knowledge by adapting its behaviour to integrate more seamlessly with the user's environment. The application, for instance, could automatically re-route a visually impaired traveller to avert hazardous excavation work on a pavement.
- *Contextual resource discovery*. Application discovers other resources within the same context as itself and exploits these resources while the user remains in the same context. For instance, when people travel abroad, local mobile phone networks often send text messages to inform people of where they can find useful information about the immediate environment. A context-aware device could automatically acquire this information on behalf of the user, such as downloading train timetables.
- *Contextual augmentation*. Application allows the user to augment the environment with information at a specific location for other users. For instance, a visually impaired traveller could leave a location-specific electronic message about a hazard. Other visually impaired travellers would only receive this message if the same location were visited.

As mentioned earlier, existing GPS-based systems for visually impaired people offer limited contextual detail since most are based only upon contextual sensing. Context-aware computing could therefore extends the visually impaired person's spatial awareness and knowledge, and provides more support, flexibility and opportunities for independent mobility involving a wide-range of activities.

7.4.4 Application of Context-aware Principles

The purpose of this section is to illustrate, using three case studies, the application of context-awareness in research and development. Currently, only one prototype that explicitly draws upon context-awareness has been developed to assist the distant navigation requirements of visually impaired people. For this reason, we will describe two other prototypes, namely, a context-aware mobile tourist guide, and a location messaging system.

Case study 7.1: The Drishti System

The Drishti wireless navigation system, which is being developed by researchers at University of Florida, transmits route information to visually impaired and disabled people in order to assist navigation through dynamic outdoor environments (Helal *et al.* 2001). Drishti is designed to generate optimised routes based upon user preferences, temporal constraints such as traffic congestion, and dynamic obstacles such as ongoing ground work, *e.g.* routes involving fewer hazards may be chosen over the shortest route. Re-routing is supported if unexpected events occur, and travellers can also add notes about certain conditions or problems encountered as a reminder if the route is revisited (such as a pothole on the road).

Drishti integrates several hardware components including a wearable computer, an integrated GPS/Beacon/Satellite receiver, an electronic compass, and components for various wireless networks. Software components include a spatial database called ArcSDE to manage GIS datasets; a route store called NetEngine to define, traverse and analyse complex geographic networks; and a mapserver called ArcIMS to serve the GIS datasets over the Internet. Communication between the user and interface is accomplished using ViaVoice which transmits detailed explanatory cues using text-speech software and also executes verbal commands of the user using voice recognition.

While the Drishti system provides many useful levels of functionality, the system is limited to outdoor environments and was tested using a very precise task and scenario, *i.e.* through a university campus. Testing this system in more contexts, such as indoor, and for other mobile tasks and visually impaired user groups would be more challenging and would involve acquiring far more contextual information; a technological issue discussed in Section 7.4.5. Other capabilities of context-awareness could also be explored such as allowing visually impaired travellers to disseminate notes to others, *e.g.* warning others of potholes in the environment. This capability is discussed in case study 7.3.

Case study 7.2: Tourist information systems

Mobile information systems for tourists have also been an application of context-aware computing. For this case study, we have chosen to discuss the GUIDE prototype, which is being developed at the Distributed Multimedia Research Group from Lancaster University (Cheverst *et al.* 2000).

GUIDE is a mobile tourist guide that transmits geographical information regarding points of interest that has been tailored to suit the tourist's personal and environmental contexts. The context-sensitive application supports different activities of the tourist which include:

- Acquiring context information regarding the current location.
- Navigating the city or local area using a map.
- Generating tours of the city and then navigating to landmarks—the application optimises tours based on opening and closing times, the likely busyness of the attraction, the distance between attractions, and the most aesthetic route.
- Communicating with other visitors or the tourist information centre by sending a text message.

An example of the GUIDE prototype is illustrated in Figure 7.3.

The hand-held GUIDE system operates on a Fujitsu TeamPad 7600 device, and its software was designed similar to an Internet browser for immediate familiarity. A cell-based wireless communication infrastructure is used to broadcast both location and dynamic information to mobile GUIDE units. Location positioning is determined from location messages that are received from strategically positioned base stations.

The human issues associated to the design of GUIDE, or tourist information systems in general, concern the extent to which they can be 'picked up and used'

Figure 7.3. The GUIDE prototype (reproduced with kind permission of Dr Keith Cheverst)

by people from different demographic and cultural backgrounds. Tourists visiting a city would have contrasting expectations of supported activities, levels of functionality, and methods of user-interface interaction. The versatility of the system is also critical, as there are likely to be (i) vast differences in how tourists would wish to experience a city, and (ii) situations where tourists make unpredictable/incidental decisions regarding future activities. Additionally, the tourist's dependency on the system would need to be measured against actual city experience.

Case study 7.3: Location messaging systems

Both previous case studies concern systems that transmit localised and personalised information to the user, either prior to a tour/journey, or as a route is experienced. The third case study, however, describes a system which is centred on an entirely different capability of context-aware computing; that of contextual augmentation (described in Section 7.4.3).

The stick-e document, described by Brown (1996), is a framework for representing this capability in the design of a wide variety of context-aware applications. It is described how the user, who would possess a location-aware mobile device, can leave stick-e notes (which are the electronic equivalent of a Post-It™ note) at specific locations in space. When the user revisits this location, the message is triggered and displayed to the user. These notes can also be broadcast and exchanged to a wider audience, if this is considered to be a desirable design feature. In this example only the location is used to determine when a note is triggered, though Brown also describes how other elements of context can be included to make the presentation of notes more useful. Triggering events could be dependent on (i) the adjacency of other objects or people, *e.g.* when the user is in the presence of a particular friend, (ii) critical states, such as temperature thresholds, and (iii) time ranges, *e.g.* only display a note about a tourist attraction during opening hours.

The exchange of notes between users would not normally be done singularly, *i.e.* one note at a time; rather, a stick-e document would be exchanged containing a group of notes that have a logical relationship, such as notes relating to a particular activity or tourist trail. The dissemination of documents could be (i) downloaded to the mobile device using the Internet, (ii) exchanged between two devices using beaming, such as Bluetooth, or (iii) loaded into the mobile device *via* a wireless link (in this method, notes within the document could be updated regularly in order to reflect environmental change).

The stick-e document would have obvious benefits for communities of visually impaired people, who could share experiences about different situations and environments, *e.g.* alerting others of hazards. The management, dissemination, and presentation of documents and notes are critical technological, human and social issues. Large amounts of notes would need to be managed, some of which may be irrelevant, inappropriate, or out of date. Adequate controls would need to be put in place in order to ensure notes are not displayed repeatedly if the user decides to revisit a location for the second or third time. Much of the triggering states may have to be customised by the user, requiring considerable time and knowledge of how the system is likely to respond. Lastly, sharing notes with others has the potential to get obtrusive and tedious, especially where there are conflicts of interest.

7.4.5 Technological Challenges and Unresolved Usability Issues

For the capabilities of context-awareness to be realised, context-aware design needs to possess efficient and effective software processes for acquiring, storing, interpreting, and utilising contextual information, in order to make accurate inferences regarding a user's task and situation. This is a complex and challenging process due to the following reasons:

- *Context is acquired from sensors*. Many heterogeneous devices need to communicate with one another, possibly using different technologies, components, and software.
- *Context is dynamic and multidimensional.* The user and environment constantly evolve, *e.g.* within the environment new buildings get erected, fashions and trends change; a user may change their preferences/dislikes; *etc.*
- *Data needs to be abstracted*. Interpreting contextual information can be extremely complicated since different varieties of contextual information need to be combined.

In addition to the technological issues, context-aware designers need to address many unresolved usability issues, which will ultimately determine whether these devices are to seamlessly integrate into people's lifestyles. Some of these issues are also particularly important and applicable to visually impaired people, which will now be described.

Information push vs information pull. Cheverst *et al.* (2000) differentiated between two contrasting styles of how a user may wish to acquire context-aware information from an application. Information *push* is where the actual presentation of information is triggered by contextual events, *e.g.* a location change, whereas information *pull* is where the user decides when information is presented. Since the timing of contextual information is critical to visually impaired people when negotiating environmental hazards, there may be situations when one style is more appropriate than the other, *e.g.* information push may not be suitable when a visually impaired person crosses a busy pedestrian crossing. A prioritisation of current activities against the cognitive benefit of new contextual information for different tasks, situations, and environments needs to be the topic of further research.

Privacy and security issues. One of the technological motivations of context-awareness is to create a global network of computers, servers, and databases in order to facilitate the exchange of detailed contextual information. As a consequence, appropriate filters must be integrated to ensure that the perceived level of privacy and security experienced by users is managed effectively. It is likely that there will be discrepancies in how accepting different users are of having their location and identity disclosed to service providers, friends, family, or work colleagues, especially for different tasks and situations, *e.g.* at work *vs* in the pub. Of course, some people may prefer no one to know their location.

Multimodality issues associated to multimodal (using multiple technologies) output. Based upon the user's task and situation, certain output technologies may be more appropriate. Speech output, for instance, may be better for visually demanding tasks, whereas a visual display may be better for illustrating spatial relations. Also, contextual information may be better transmitted when spanned across more than one sensory channel, such as through hearing and touch. Ross and Blasch (2000) argue that the best interface for visually impaired people is a combination of speech output and tactile cues using a tapping interface. More research is needed, though, to understand (i) which types of contextual information should be represented using different forms of output technologies, and (ii) individual user preferences, especially for people with different visual impairments (Bradley and Dunlop 2003).

Social implications. One of the ideas behind context-aware computing is to facilitate groups of people in sharing contextual information. There are, however, some important social implications surrounding this notion, with regards to how this is information is being shared and distributed to others (Bellotti and Edwards 2001). What rights should be given to other people when they send information of potential interest to the user? For instance, the user's perception of what is useful and important may be entirely different to the sender. Ideally, context-aware computing research needs to address how different social relationships and situations affect how contextual information is shared and disseminated.

7.5 Test Procedures

The design and evaluation of electronic navigation aids brings together two very different research domains:

- Cognitive mapping research has a long history of investigating peoples' understanding of the geographical environment they live in and how they build cognitive maps of that world.
- Since the seventies, human computer interaction (HCI) researchers have been developing methodologies and approaches to the design, implementation and evaluation of easier-to-use interactive computer systems.

This section will give a very brief overview of these approaches, but will focus mainly on HCI approaches and their application to the design and evaluation of electronic navigation aids.

7.5.1 Human Computer Interaction (HCI)

User-centred design is the core to much good practice in human-computer interaction and much of modern software engineering. This approach to interactive system design puts users at the heart of the design processes for the systems that they will use. Having users involved in the design and development phases of applications, particularly at the start, overcomes the difficulties in traditional requirements analysis of modelling users' experience, understanding, working practices and desires. User-centred design has been proven to not only lead to more usable systems but to save development time in large projects. Building on the premise that we often cannot describe what we want but can critique what we have, user-centred design processes often make heavy use of prototypes to aid the dialogue between designers and end-users.

Prototypes are partial implementations of the system that can be used to engage users and to run initial usability trials by getting users to perform tasks with the prototype system. These prototypes can be of varying quality and completeness: in the early stages of product development simple hand drawn paper prototypes are used as the basis for discussions with users about their requirements. When designs are more developed, semi-functional systems can be implemented—these often lack features and the robustness of the final system but still give an impression of using that system. So-called *Wizard of Oz* experiments can be carried out using completely impersonated system functionality—that functionality being provided by a human acting as if they were the system. For example, an audio-based navigation system could be impersonated by a designer by simply reading navigation instructions to a user wearing a hands-free mobile phone headset. The Wizard of Oz technique can be very useful where considerable development is needed to assess the value of a real system.

If the intended users of a system cannot be central to the design team, one solution is often to substitute them with *personas*, shown in Figure 7.4. A persona is a description of a fictitious intended user of the system. Personas are given names and often embellished with personal information to make the fictitious person feel like a real person to the design team.

Evaluation of how people perform with systems is fundamental to user-centred design. Traditionally this evaluation was of desktop computer systems intended for office use; as such, the evaluation focussed heavily on users performing tasks sitting at a computer in an office environment. Many of the traditional techniques have been adapted for use in evaluating more modern systems, such as mobile computing devices (including palmtop computers, Personal digital assistants and mobile phones) and context-aware systems. There are too many techniques for interactive system evaluation to discuss here and they are well documented in HCI text books (*e.g.* Dix *et al.* 2003; Preece *et al.* 2002) and usability evaluation texts (*e.g.* Rubin 1994). Two traditional interactive evaluation techniques, *think-alouds*

Stuart is a 23-year-old Glaswegian who lost his sight in an accident as a child. He is studying computing at Strathclyde University and hopes to graduate next year with a 2:1 honours degree. He has lived in Glasgow all his life but has travelled fairly extensively in Europe on short holidays with friends and family. He is considering moving away from home for his first job and is currently applying for jobs in Edinburgh and London. Stuart follows Partick Thistle football club, is heavily into the mixing/DJ scene in Glasgow but, currently, doesn't have a partner (solving one dilemma over moving city).

Figure 7.4. Sample persona

and *controlled experiments*, and one evaluation technique used in HCI exclusively for mobile system evaluation, *percentage preferred walking pace*, are worthy of further discussion here.

A *think-aloud* is a very simple, cost-effective, evaluation approach that is both easy to run and can produce significant feedback to designers in a very short time. To run a think-aloud the designer needs:

1. An interactive system to test
2. Normal support materials for first-time system use, such as manuals, tutorials *etc.*
3. A user (one per think-aloud session—normally a trial will compose several sessions)
4. A list of tasks for the user to do with the system

The user is given the system and asked to work through the task sheet "thinking aloud", *i.e.* they are asked to verbalise their thoughts as they work through the set of tasks using the system. In an office environment, the experimenter sits behind the user taking notes and occasionally giving neutral prompts such as "what are you thinking now?". A video recording is normally also made for later analysis—ideally this captures multiple camera angles covering the user's face, hands and the screen plus audio from the user and the system. The video can also be used after the formal session to discuss design problems with the user and engage him/her in possible redesigns.

When well designed, the approach provides considerable scope for users to act unexpectedly and use the system in ways that the designers never intended, never mind tested. Despite being designed for desktop office applications, think-alouds can easily be adapted to mobile devices such as computer-based navigation aids and context-aware systems. The only problem is recording the events, but the experimenter can still learn many lessons by simply noting them down in a notepad as they walk behind the user.

While valuable lessons can be learned by designers on how to make their system more usable using a few think-aloud sessions, concrete numbers and comparisons are hard to gain from this approach: the very nature of think-alouds results in them impacting on most numeric measures that could be used for concrete comparison, *e.g.* the time it takes a user to complete the tasks. Controlled experiments complement think-alouds where either a research hypothesis is being tested or two systems are being compared. Controlled experiments are often run in a similar environment to think-alouds but with more focussed tasks, no interaction between users and experimenter once the study is underway and no need for the user to verbalise their thoughts. After an initial training period, the experimenter stays silent and simply records the user's behaviour—most often focussing on numeric information such as the time to complete the tasks and the number of task-related errors the user makes in the process. When correctly designed, this form of experiment leads to statistically testable numeric results, thus it is important for scientific studies or system comparisons of the form *system A is better/faster/less-error-prone/safer than system B*. However, such experiments are often less helpful in finding design problems than think-alouds: the environment can be so constrained that unexpected user behaviour is very rare. Controlled experiments must reduce as many environmental variations as possible in order to get statistically significant results with a reasonable number of users. It is thus difficult, but not impossible, to run valid controlled experiments on mobile devices in mobile settings: for example, the effects of, say, traffic conditions are likely to swamp any difference between two systems when measuring how long it takes a person to navigate across a city. The variation in conditions has to be taken into account in the experimental design and, often, many more users are required to achieve statistically significant results. As a compromise between realistic settings and the desire for solid numerical results, some researchers have carried out trials in semi-constrained environments, for example Brewster (2002) asked users to complete tasks on a hand-held computer while walking laps of a quiet outdoor route.

Adults tend to have a preferred walking speed (PWS) that they will walk at if allowed to walk freely. A person's walking pace is, however, often affected by their concentration on other activities, such as navigating in an unknown environment or walking while using a mobile phone or a navigation system. This reduced speed can be expressed as a percentage of the user's preferred walking speed (PPWS) and can be used as an estimate of how much, say, the mobile computing device is interfering with the user's walking. An ideal navigation system would allow the user to walk at their particular PWS. If this pace is known, then real systems can be compared on what the PPWS is while using the systems. This approach has been used both to test hand-held computers (*e.g.* Brewster *et al.* 2003) and navigation systems for visually impaired people (*e.g.* Petrie *et al.* 1998; Soong *et al.* 2000).

7.5.2 Cognitive Mapping

Kitchin and Blades (2002) provide a comprehensive literature study of cognitive mapping methods and categorise methods for individual studies as follows.

- *Unidimensional data generation.* Studies in this category involve studying cognitive models a single dimension at a time, focusing either on distance or angle. Various approaches to distance analysis can be used that attempt to overcome different problems in getting people to express what their internal model is. These include simple magnitude estimates, *e.g.* "if Glasgow–London is 100 units, how far is Glasgow–Dublin?" and rating distances into different categories, *e.g.* very near, near, medium, far, very far, *etc.* Direction estimates usually involve either standing (or imagining oneself standing) at a location and pointing to another or drawing the direction from one location to another on paper.
- *Two-dimensional data generation.* Simple graphic approaches to 2D studies involve getting the experimental subject to draw maps of an area. Unfortunately, these drawings can be affected not only by the subject's cognitive map but by their ability to express that through drawing—Kitchin and Blades discuss alternatives to try to reduce this problem. Completion tasks are one solution where subjects complete a partial map—either free hand drawing additional information on pre-prepared maps or filling in blank spaces on the map.

- *Recognition tasks.* Rather than relying on subjects descriptive abilities, recognition tasks simplify the subject's task by asking them to select the correct map or map segment from a multiple-choice selection where the incorrect choices are variants of the correct map, *e.g.* skewed or rotated.
- *Qualitative approaches to studying cognitive maps.* The above methods are either inherently quantitative in nature or can be easily analysed quantitatively. As with HCI, cognitive mapping researchers have made use of think-aloud protocols among other techniques, to study, for example, how people learn spatial information from maps.

Kitchin and Blades also discuss other studies—particularly ones that attempt to verify models or test subjects' cognitive maps, *e.g.* comparing subjects ability to follow a route after previously following the route either in real-life or on video.

7.5.3 Overall Approach

The goals of much of human-computer interaction practice differs considerably from that of both human-computer interaction research and cognitive mapping research: both research fields are intent on understanding people whereas HCI practice is about designing systems that are easy to use. This leads to a natural focus reflected here on predominantly qualitative methods for HCI practitioners wanting to discover usability problems and quantitative methods for analysing and understanding people's mental models.

Central to modern interactive systems design is the need for multidisciplinary teams. For the successful design of navigation systems having team members who understand cognitive mapping research and methods (in our case, particularly the work on cognitive mapping of visually impaired people) and team members who understand HCI approaches is imperative to successfully designing mobile navigation aids.

7.6 Future Positioning Technologies

Location-aware services (the awareness is limited to just location information) are the most widespread context-aware services, however even these are still in their infancy. Focussed mainly around research systems, location-aware services are starting to become more mainstream: initially driven by emergency service legislation, discussed below, and some tourism applications (*e.g.* Cheverst *et al.* 2000). As these new technologies are more widely adopted, the infrastructure will be put in place and equipment prices will drop enabling more widespread and disparate location-aware services. The key to success of many of these new services, however, is improving both the "at best" accuracy of location information and the reliability of that information in normal use. Providing navigational assistance to blind pedestrians, for example, is a major test for both accuracy and reliability, in which it is necessary to consistently know the user's location to within about 2 m to provide independent navigational advice. This section provides a brief overview of some of the technological developments that might make such high accuracy location information possible, much of it is covered in more detail by the excellent reports from the Finish Personal Navigation (NAVI) programme (*e.g.* Kaasinen *et al.* 2001).

Mobile phone base stations have a geographic area, or cell, of coverage—phone calls to a mobile are directed to a given base station based on the mobile phone being last detected within a geographic cell allocated to that base station (see Figure 7.5). Approximate location-aware services can be provided by simply providing access to the cell information, *e.g.* a single taxi number could be provided that will be redirected to a local taxi company for the cell the mobile phone is currently, in anywhere in that country. Mobile phone cells, however, vary greatly in size: in busy urban areas, where cells are smallest and closest together, using cell information can locate a user to within 200 m. However, in rural areas the accuracy can be as low as 30 km and might not even locate the user on the correct island or country as signals travel well over water.

In the U.S.A., emergency mobile phone legislation (see http://www.fcc.gov/911/enhanced/) required, by the end of 2005, that 95% of mobile phones were capable of giving an accurate location when an emergency (E911) call is made. (Note that similar initiatives are under way for 112 calls within the E.U.) The E911 legislation requires that the location must be provided to within 50 m for 67% of calls and 150 m for 95% of calls from location-enabled handsets. This legislation has led to considerable research into accurately providing location information at this level of detail. There are two basic approaches: make improved use of mobile phone networks to locate phones more precisely than simple cell information or augment phones with GPS technology. Both these approaches make use of existing infrastructure, *e.g.* base stations or satellites, reducing the costs of deployment, but add computational load to the base stations, the handsets or both. Once appropriate legal and technical frameworks are in place to protect privacy, it is likely to roll out to other services.

Mobile phone base stations can measure the transmission delay of a signal to/from a handset and the angle at which the signal is received at the base sta-

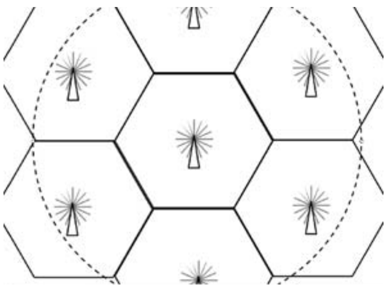


Figure 7.5. Simplified mobile phone cell model with each *hexagon* representing a cell area and the *circle* representing the actual area of signal coverage of centre cell

tion. In urban areas a phone may be allocated to a given base station for phone calls but often "sees" the signal from base stations in neighbouring cells or of other companies. Trials have shown that using triangulation of signal strengths or angle of signal receipt can achieve the E911 requirements within urban areas and locate users to within about 45 m. However, while the targeted 45 m accuracy is useful for emergency services and some mobile services it is too poor for navigation.

Current global positioning systems can provide accuracy of around 10 m, which is suitable for distant environment navigation. However, because of the need for line-of-sight to satellites, GPS does not work indoors and receivers often have trouble in high-rise cities (where the buildings block sight of satellites and can cause confusing reflections of the signals). Furthermore, GPS receivers can take several minutes to calculate a location when first switched on—too long for emergency calls. Assisted-GPS technology allows higher accuracy than traditional GPS with less computation on the handset through the use of base-stations that know the paths of satellites over the near future and have clear view of the skies. Assisted-GPS enabled mobile phones use the mobile network to get additional satellite information together with modified, and simpler, GPS hardware. This combination can consistently provide sub-10 m accuracy outdoors in cities and 40 m accuracy indoors. In the long-term, and complementing GPS, the European Union Galileo system is expected to have 30 satellites in place by 2008 improving the accuracy of GPS to about 1 m.

Independently of the mobile phone network, many companies within business areas of cities now have wireless ethernet (a.k.a. Wi-Fi or 802.11) installed to provide their employees with high speed wireless networking. If the location and identity of base stations are known in advance, this information can be used together with signal-strength, to provide highly accurate, *e.g.* sub-1 m, positioning within buildings equipped with wireless ethernet. While these networks are often private, the signals and identification information often leaks outside buildings into streets and public spaces. There are some initiatives looking at using these signals to provide location-aware services (*e.g.* Small *et al.* 2000) and the approach nicely complements GPS in busy urban areas due to the lack of sky visibility.

In the relatively near future, combinations of these technologies are likely to provide ubiquitous, relative cheap, and high accuracy location information that can be exploited for many mobile applications; in particular achieving the accuracy and reliability requirements for assisting visually impaired navigation.

7.7 Chapter Summary

The terms 'orientation', 'mobility' and 'navigation' were initially defined and discussed with respect to visually impaired people. Commonly used mobility aids, and the limitations of their use during navigation, are then examined. In order to investigate those limitations more deeply, the next section describes the area of cognitive mapping which refers to the process by which individuals acquire, learn, develop, think about and store data relating to the everyday geographic environment, such as locations, attributes, and landmark orientations to navigate. This provided an insight into the cognitive maps of visually impaired people and how they are used to orientate and navigate. Information regarding the distant environment, which cannot be acquired from commonly used mobility aids, was found to be critical for independent mobility. This led to a review of technologies designed to support distant navigation, such as the use of GPS systems for navigating outdoors. Technologies used to transmit navigational information were also discussed. In the next section the area of context-aware computing is introduced as a way to combine technologies in order to deliver more useful and relevant information. While the capabilities of context-awareness offer a huge potential, there still remain unresolved technological and human issues, particularly the security and privacy of information regarding the user's location. In order to illustrate how such human issues would be investigated, the next section describes test procedures involving end-users, the results of which are used to improve usability. A technological requirement, on the other hand, is to improve the accuracy and reliability of location-aware systems for assisting visually impaired navigation. This is the subject of the last section which describes how this might be achieved in light of recent advances in positioning technology.

7.7.1 Conclusions

This chapter initially introduced the principle of independent mobility, where visually impaired people are able to travel freely through the environment without feeling restricted to known routes or destinations. While traditional aids provide an important tool for the identification of obstacles in the immediate environment, these aids need to be complemented with systems that provide information regarding the distant environment if independent mobility is to be realised. Although location-aware systems have provided a valuable first step towards this principle, these systems need to be expanded in order to include more contextual detail of the traveller's environment. Various methods and techniques, which are found in HCI and cognitive mapping research, can provide invaluable input to application developers who need to establish which user behaviours to support and what contextual information would be of use. Context-aware computing moves closer to this principle, where applications are able to sense, store, interpret, and manage large amounts of contextual detail in order to infer the user's situation and activity. While future improvements in location precision technology and other technological advancements are paramount, the biggest hurdle perhaps will be the extent to which the human issues are addressed.

This is the key focus of our research and much of the user-centred ubiquitous community. We work closely with various organisations representing the needs of visually impaired people, and use them to recruit participants for our user studies. Our most recent study, discussed by Bradley and Dunlop (2004) involved testing a mock-up application that adjusts contextual information for:

- People with different visual impairments, *i.e.* three groups consisting of people who have a central vision loss, a peripheral vision loss, or who are registered blind.
- Different contextual environments, *i.e.* indoor *vs* outdoor.

Significant differences exist between these groups, and between indoor and outdoor environments. Our overall goal is to develop a design framework that will help application developers to include pertinent human issues in design in order to build more useful and relevant mobile context-aware services for visually impaired people.

Questions

Q.1 What are the limitations of traditional mobility aids?

Q.2 Describe the three main theories of how spatial information is encoded and discuss the factors that may influence this process.

Q.3 Which technologies can be used to provide distant navigation support to visually impaired people within indoor and outdoor environments? What are the constraints of each?

Q.4 What are the limitations of traditional user interfaces, and how could those limitations be addressed by context-aware computing?

Q.5 Describe how a context-aware application might infer a user's task or situation. Illustrate your answer with a scenario.

Q.6 What are the capabilities of context-aware computing, and how have these capabilities been applied in practice?

Q.7 Discuss the issues surrounding how contextual information is acquired, stored, interpreted, and utilised. What are the design issues with respect to the user?

Q.8 What types of human-computer interaction approaches and cognitive mapping methods exist? Describe the benefits and limitations of each.

Q.9 Discuss future technologies and techniques for providing more accurate locations.

Q.10 How would context-aware services be facilitated through the ideology of ubiquitous or pervasive computing? Illustrate with examples.

Q.11 Consider, or discuss in groups, your views on the safeguards that would need to be in place before you would use a device that constantly knows your location to within 5 m.

Projects

P.1 Current technologies enable the following information to be sensed about the user:

Type of technology	Contextual information about the user
GPS, Active badge, mobile phone cells	Indoor and outdoor location of the user
RF beacons	Transmission of location-specific audio information using RF
Bluetooth and infrared	Detection of surrounding objects, *e.g.* printers, other mobile devices
Accelerometers sensors	Speed of movement
Electronic compass	Orientation of body, *e.g.* upright or lying down
Temperature sensors	Outside temperature
Physiological sensors	Heart rate, body temperature, tone of voice, *etc.*
Personal diaries	Social and business arrangements, activities, *etc.*
Web-based servers	Weather forecasts, train and bus timetables, *etc.*
Camera technology and image processing	Face recognition to detect surrounding people

Describe a scenario that would be encountered by a visually impaired traveller. Investigate the user activities and requirements within this scenario, and then explore which types of contextual information could be sensed by the application to either infer or support these activities. Consider the capabilities of context-aware computing, and the technological and usability issues that may arise.

P.2 Your task is to design and test two different styles of communicating navigation information to a visually impaired traveller. By programming a Palmtop, or using a minidisk/MP3 player, to play audio files along a pre-determined route, investigate the issues surrounding information push *vs* information pull (discussed in Section 7.4.5) using a 'Wizard of Oz' style experiment.

References

Allen, G.L., 1999, Spatial abilities, cognitive maps, and way-finding: Bases for individual differences in spatial cognition and behaviour, In Golledge, R.G. (eds), Wayfinding Behavior, Cognitive Mapping and Other Spatial Processes, John Hopkins, Baltimore, pp. 301–321

Bellotti, V., and Edwards, K., 2001, Intelligibility and accountability: Human considerations in context-aware systems, J. Human-Computer Interaction. Vol. 16, pp. 193–212.

Bigelow, A.E., 1996, Blind and sighted children's spatial knowledge of their home environments, International Journal of Behavioural Development, Vol. 19, pp. 797–816

Bradley, N.A., and Dunlop, M.D, 2003, A pathway to independence: way-finding systems which adapt to a visually impaired person's context, Procs Symposium on Assistive Technology – Towards a better life, Glasgow, 30 April 2003, pp. 23–27

Bradley, N.A., and Dunlop, M.D., 2004, Investigating design issues of context-aware mobile guides for people with visual impairments, Procs Workshop on HCI in Mobile Guides, Mobile HCI04, Glasgow, UK, 13th September 2004

Brewster, S.A., 2002, Overcoming the lack of screen space on mobile computers, Personal and Ubiquitous Computing, Vol. 6, No. 3, pp. 188–205

Brewster, S.A., Lumsden, J., Bell, M., Hall, M., and Tasker, S., 2003, Multimodal 'Eyes-Free' Interaction Techniques for Wearable Devices, Procs ACM CHI 2003 (Fort Lauderdale, FL), ACM Press, Addison-Wesley, pp. 463–480

Brown, P.J., 1996, The Stick-e Document: a Framework for Creating Context-Aware Applications, Procs EP'96, Palo Alto, pp. 259–272

Bunt, H., 1997, Context and Dialogue Control, Procs 1st International and Interdisciplinary Conference on Modeling and Using Context, CONTEXT 1997, Rio de Aliceiro, Brazil, pp. 130–149

Cheverst, K., Davies, N., Mitchell, K., Friday, A., and Efstratiou, C., 2000, Developing a Context-Aware Electronic Tourist Guide: Some Issues and Experiences, Procs CHI 2000, 1–6 April 2000, pp. 17–24

Clark-Carter, D.D., Heyes, A.D., and Howarth, C.I., 1986, The efficiency of walking speed of visually impaired pedestrians, Ergonomics, Vol. 29, pp. 779–789

Cornell, E.H., and Hay, D.H., 1984, Children's acquisition of a route via different media, Environment and Behaviour, Vol. 16, pp. 627–64

Dey, A.K., and Abowd, G.D., 2000a, Towards a better understanding of context and context-awareness, Procs CHI 2000 Workshop on The What, Who, Where, When, and How of Context-Awareness, The Hague, Netherlands, April 2000

Dey, A.K., and Abowd, G.D., 2000b, The Context Toolkit: Aiding the Development of Context-Aware Applications, Workshop on Software Engineering for Wearable and Pervasive Computing, Limerick, Ireland, June 6, 2000

Dix, A., Finlay, J., Abowd, G.D., and Beale, R., 2003, Human Computer Interaction, Prentice Hall

Downs, R.M., and Stea, D., 1997, Cognitive Maps and spatial behaviour: Process and Products, In R.M. Downs and D. Stea (Eds), Image and Environment, Chicago, Aldine, pp. 8–26

Farmer, L., 1980, Mobility devices, In R.L. Welsh and B.B. Balssch (Eds), Foundations of Orientation and Mobility, American Foundation for the Blind, New York, pp. 357–412.

Franklin, N., 1995, Language as a means of constructing and conveying cognitive maps, In Portugali, J. (Ed), The construction of cognitive maps, Kluwer Academic Publishers, pp. 275–295

Gärling, T., Book, A., Lindberg, E., and Nilsson, T., 1981, Memory for the spatial layout of the everyday physical environment: Factors affecting rate of acquisition, J. Environmental Psychology, Vol. 1, pp. 263–277

Gibson, J.J., 1979, The ecological approach to visual perception, Houghton Mifflin, Boston

Golding, A.R., and Lesh, N., 1999, Indoor navigation using a diverse set of cheap, wearable sensors, Procs Third International Symposium on Wearable computers, San Francisco, CA. pp. 29–36

Golledge, R.G., 1978, Representing, interpreting and using cognized environments, Papers and Procs Regional Science Association, Vol. 41, pp. 169–204

Golledge, R.G., and Stimson, R.J., 1987, Analytical Behavioural Geography, Croom Helm, London

Golledge, R.G., and Stimson, R.J., 1997, Spatial Behavior: A Geographic Perspective, Guildford Press, New York

Haken, H., and Portugali, J., 1996, Synergetics, inter-representational networks and cognitive maps, In Portugali, J. (Ed.) The construction of Cognitive Maps, Kluwer, Dordrecht, pp. 45–67

Helal, A.S., Moore, S.E., and Ramachandran, B., 2001, Drishti: An integrated Navigation System for Visually Impaired and Disabled, Procs 5th International Symposium on Wearable Computers, Zurich, Switzerland

Hill, E., and Ponder, P., 1976, Orientation and Mobility Techniques: A Guide for the Practitioner, American Foundation for the Blind, New York, p. 114

Holland, S., and Morse, D.R., 2001, Audio GPS: spatial audio in a minimal attention interface, Procs Third International Workshop on Human Computer Interaction with Mobile Devices, IHM-HCI 2001, Lille, France, September 2001

Jonsson, E., 2002, Inner Navigation: why we get lost and how we find our way, Scribner, New York, pp. 27–126

Kaasinen, E., Ikonen, V., Ahonen, A., Anttila, V., Kulju, M., Luoma, J., and Södergård, R., 2001, Products and Services for Personal Navigation – Classification from the User's Point of View, Report of Personal Navigation (NAVI) programme, Finnish Ministry of Transport and Communications, http://www.vtt.fi/virtual/navi/

Kemmerling, M., and Schliepkorte, H., 1998, An Orientation and Information System for Blind People based on RF-Speech-Beacons, Procs TIDE, Helsinki, July 1998, http://www.stakes.fi/tidecong/651kemm.htm

Kitchin, R.M., and Blades, M., 2002, The Cognition of Geographic Space, I.B. Tauris, London, pp. 33–57

Kitchin, R.M., and Jacobson, D., 1997, Techniques to collect and analyze the cognitive map knowledge of people with visual impairments or blindness: Issues of validity, Journal of Visual Impairment and Blindness, July-August, pp. 360–376

Kitchin, R.M., Blades, M., and Golledge, R.G., 1997, Understanding spatial concepts at the geographic scale without the use of vision, Progress in Human Geography, Vol. 21, No. 2, pp. 225–242

LaGrow, S., and Weessies, M., 1994, Orientation and Mobility: Techniques for Independence, Dunmore Press Ltd., New Zealand, p. 9

Loomis, J.M., Golledge, R.G., and Klatzky, R.L., 2001, GPS-based navigation systems for the visually impaired, In W. Barfield and T. Caudell (Eds.), Fundamentals of Wearable Computers and Augmented Reality, Lawrence Erlbaum Associates, Mahwah, NJ, pp. 429–446

Makino, H., Ishii, I., and Nakashizuka, M., 1997, Development of navigation system for the blind using GPS and mobile phone combination, Procs 18th Annual International Conf. IEEE Engineering in Medicine and Biology Society, Vol. 18, Pts 1–5

Moran, T., and Dourish, P., 2001, Introduction to Context-Aware Computing, Special Issue of Human-Computer Interaction, Vol. 16

Pascoe, J., 1998, Adding Generic contextual capabilities to wearable Computers, Procs 2nd International Symposium on Wearable Computers, pp. 92–99

Petrie, H., 1995, User requirements for a GPS-based travel aid for blind people, In J.M. Gill and H. Petrie (Eds.), Procs Conference on Orientation and Navigation Systems for Blind Persons, Hatfield, UK, 1–2 February, Royal National Institute for the Blind, London

Petrie, H., Furner, S., and Strothotte, T., 1998, Design Lifecycles and Wearable Computers for Users with Disabilities, Procs First Workshop on Human-Computer Interaction with Mobile Devices, Glasgow

Pitt, I.J., and Edwards, A.D.N., 1996, Improving the usability of speech-based interfaces for blind users, Procs 2nd Ann. ACM Conf. Assistive technologies, Vancouver, Canada, April 1996, pp. 124–130

Preece, J., Rogers, Y., and Sharp, H., 2002, Interaction Design: Beyond Human-Computer Interaction, John Wiley and Sons

Ross, D.A., and Blasch, B.B., 2000, Evaluation of orientation interfaces for wearable computers, Procs Fourth International Symposium on Wearable Computers (ISWC'00). Atlanta, Georgia, October 18–21, 2000, pp. 51–61

Rubin, J., 1994, Handbook of Usability Testing: How to Plan, Design and Conduct Effective Tests, John Wiley and Sons

Selker, T., and Burleson, W., 2000, Context-aware design and interaction in computer systems, IBM Systems Journal, Vol. 39, Nos. 3 and 4, pp. 880–891

Small, J., Smailagic, A., and Siewiorek, D.P., 2000, Determining User Location For Context Aware Computing Through the Use of a Wireless LAN Infrastructure, Institute for Complex Engineered Systems Report, Carnegie Mellon University, Pittsburgh, USA

Smith, S., 1988, Environmental Context – Dependent Memory, Chapter 2, In: Davies, G.M. Thomson, D.M. Memory in Context; Context in Memory, John Wiley and Sons, Chichester, pp. 13–34

Soong, G.P., Lovie-Kitchin, J.E., and Brown, B., 2000, Preferred walking speed for assessment of mobility performance: sighted guide versus non-sighted guide techniques, Clinical and Experimental Optometry. Vol. 83, No. 5, September, pp 243–293

Strothotte, T., Fritz, S., Michel, R., Raab, A., Petrie, H., Johnson, V., Reichert, L., and Schalt, A., 1996, Development of dialogue systems for the mobility aid for blind people: initial design and usability testing, Procs ASSETS '96, Vancouver, British Columbia, Canada, pp. 139–144

Want, R., Hopper, A., Falcão, V., and Gibbons, J., 1992, The active badge location system, ACM Trans. Information Systems, Vol. 1, pp. 91–102

Weiser, M., 1991, The computer for the 21st century, Scientific American, Vol. 265, Sept., pp. 94–104

8 Accessible Global Positioning System (GPS) and Related Orientation Technologies

Learning Objectives

Orientation has taken a significant technological leap forward in the last decade. Data of increasing accuracy from global positioning system (GPS) became available for non-military applications; computer hardware and software became more robust, portable and lightweight; WiFi and Bluetooth technology removed the need for unsightly constraining connecting wires and leads. The new generation of portable GPS orientation systems has built on these key technological developments to give visually impaired and blind people the potential tools to navigate journeys over much longer distances. For the visually impaired or blind person, even the running commentaries that these new GPS devices can give an enhanced travel experience whilst on bus-rides or a train journey. This chapter provides a personal perspective from two US engineers who were involved in creating these new systems. Various issues like the potential of the emergent technology, commercialisation issues and the use of these systems within the visually impaired and blind community are discussed. The learning objectives for the chapter include:

- Understanding the main principles of the orientation data generated by the GPS system.
- Gaining an appreciation of the issues involved in developing available advanced technology into a commercial assistive technology system.
- Understanding the different aspects of user interface design, test procedures and eventual commercialisation.

8.1 Defining the Navigation Problem

The cities we live in were not designed with visually impaired people in mind, especially when it comes to navigation. When a person who can see looks for an unfamiliar place how do they do it? They look at a map, look for visual cues like street signs, ask someone for directions, or log onto an Internet map web site. If a sighted person now puts on a blindfold and tries the same things, they will find that printed maps are useless as is looking around for street signs or other landmarks. Asking for verbal directions can be useful if someone is available but in general people do not give adequate navigational directions, because their navigational skills are visually based. For example: "Where is Big Mama's Pizza?" – "Go down five or six blocks until you come to Main Street turn left and follow this until you reach the fork in the road and take the street on the left. Then continue for about five minutes until you see the big pizza box on the side of the building. You can't miss it." Try and follow these instructions if you are blind or if you are sighted with a blindfold on.

From the illustration above, you can see that access to location information is an essential part of navigating independently. Location information is another way to describe all the environmental cues that one uses to understand one's surroundings and to navigate. Some examples are street signs, businesses and landmarks.

8.1.1 What is the Importance of Location Information?

Sighted people know when they are lost and when they are in their "*comfort zone*" based upon familiar landmarks assimilated through experience. They have mental map images based upon this information.

Consider what sighted people have to work with when they are getting around. One can easily see street signs, building names, distinct landmarks, familiar and unfamiliar settings. They can take or leave this information if and when they need it and it is a huge part of what helps them in getting around. It is not just that they can see where they are going, it is the fact that they have all this location information to guide them and to give them feedback and choices. The blind individual has a considerable disadvantage because they do not have access to location information (Loomis *et al.* 1994).

Until recently, almost all location information was print and visual and therefore not accessible to people who are blind and to most that are visually impaired. Most people who are blind or visually impaired are used to not having location information so they literally and figuratively do not usually know what they are missing. At the same time, it is fundamental to independent navigation that one uses the available tools. When detailed information is not available, blind people still make the most of the situation, however, increased location information will enhance the quality of the travel experience and will give the user more with which to work.

Another aspect of navigation for the blind, which has been under-played, is travel in vehicles. We live in an age of motor vehicles and although blind people may walk more than the average person, more than half their travel time is probably spent in a vehicle. When a blind person is in a vehicle, there is little to no contact with the signs and landmarks one is passing. It is analogous to riding in a vehicular cocoon, deprived of information even more than when walking.

8.1.2 What Mobility Tools and Traditional Maps are Available for the Blind?

The most common mobility tools used to aid in obstacle avoidance are:

- Long cane.
- Guide dog.
- Sighted guide.
- Tactile maps.

The long cane and guide dog allow the blind pedestrian to explore their environment independently. A sighted guide would make sure the blind traveller travels from point A to point B without running into obstacles. Ultrasonic devices like the Miniguide augment these other mobility tools.

In this chapter, we are focussing on electronic orientation technology, specifically that which uses the GPS, as opposed to obstacle avoidance devices for mobility like the long cane or electronic travel aids (ETAs).

When it comes to maps, blind people are not only in the dark, they are in the dark ages. Internet-based navigation, like Yahoo!® maps and MapQuest®, could be useful if the blind person has a computer screen reader but again these mapping systems are designed with one thing in mind the "sighted" mass market. These systems will try to find the most direct driving route using highways, and obeying the rules of the road. This may be a great solution if you have someone to drive you, but is useless for a blind pedestrian.

While tactile maps are great for giving an overview of a country, state, or city, they are not as good for representing detailed streets and landmarks. A physical overview plays an important role in developing a mental geographical picture. The fundamental size limitation of tactile resolution also limits its utility for quickly displaying detailed street information. A good tactile geographical overview does help to put the detailed digital map information into context if those details are accessible.

Blind travellers with ingenuity and excellent orientation and mobility skills have learnt to use their memory and other senses to access a small percentage of this location information. Specific routes are memorized including the occasional landmark to remind the blind traveller where to turn or conclude the route. It has been stated that visually impaired people have no less potential than the sighted for developing a fully integrated representation of space (Millar 1988). Notice the use of the word *potential* in that statement. Blind people can get around effectively with proper training, experience and tools.

8.2 Principles of Global Positioning Systems

8.2.1 What is the Global Positioning System?

The GPS is a worldwide radio-navigation system formed from a constellation of 24 satellites with three active spares and their ground stations (Figure 8.1). The

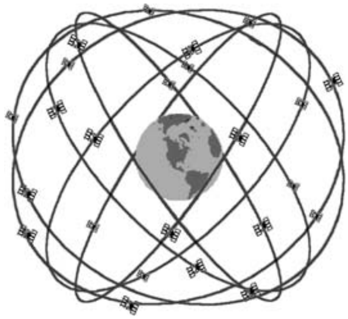


Figure 8.1. GPS satellites in orbit (image courtesy of Peter H. Dana, The Geographer's Craft Project, Department of Geography, The University of Colorado at Boulder)

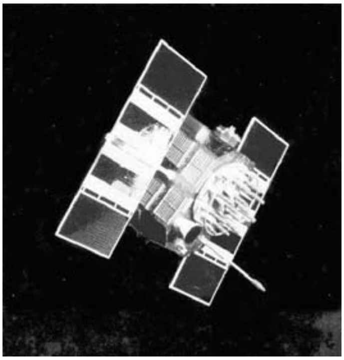


Figure 8.2. GPS satellite (image courtesy of Peter H. Dana, The Geographer's Craft Project, Department of Geography, The University of Colorado at Boulder)

satellites orbit the earth in six orbital planes at 55°, 12-h orbits, at an altitude of 20,350 km (12,644 miles) in space, weigh 862 kg (1900 lbs.), are 5.18 m (17 ft) in length, and last for about 7.5 years before they need to be replaced (Figure 8.2).

The ground stations (also known as the "Control Segment") verify that the GPS satellites are functioning properly and keep track of their exact position in space. If there are any discrepancies between the satellites and the ground stations, the master ground station will transmit the corrections to the satellites themselves. This will ensure that the GPS receivers have the correct data from the satellites (TNL 2001). The concept behind GPS is to use satellites in space as reference points for locations here on earth. Picture yourself as a point on earth with the satellites circling above you in space. The GPS software measures the distance from the satellites to your GPS receiver. This is calculated by measuring the time it takes for the signal to travel from the satellite to the receiver. By gathering data from at least three satellites (also called trilateration; Figure 8.3), the GPS receiver is able to calculate your position on earth and sends the latitude and longitude coordinates to whatever device (including a computer) that can make use of them.

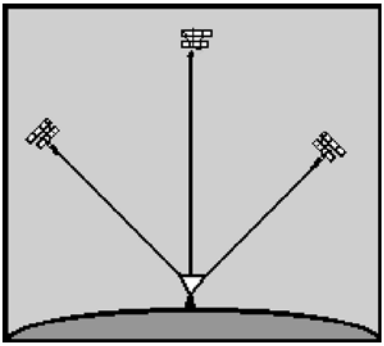


Figure 8.3. Illustration for the trilateration of GPS satellites

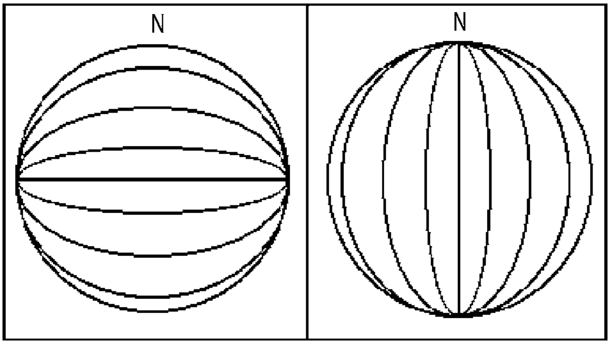


Figure 8.4. Illustration for latitude (*left*) and longitude (*right*)

Latitude lines run horizontally across the globe and longitude run vertically. Every point on the planet has a latitude and longitude and the GPS means this point can be given a meaningful name (Figure 8.4).

8.2.2 Accuracy of GPS: Some General Issues

On average, commercial GPS receivers are accurate within 9.15 m (30 feet). So, instead of viewing GPS position as a pinpoint, consider it as a bubble of radius 9.15 m around your position (Figure 8.5). However, various general factors can make the satellite information more or less accurate.

For a GPS receiver to work properly, it needs to have a clear view of the satellites. That means that GPS receivers do not work in places where all the satellites can be blocked, such as indoors, in tunnels, or in subways, for example. Poor GPS reception is also known to occur on streets in big cities when some satellites are screened by skyscrapers (also called urban canyons), and in areas surrounded by very tall mountains or forests.

Another situation where a GPS receiver might not be as accurate would be when all the satellites are coming from the same direction. Due to the significant distance between the satellites and the GPS receiver, the satellites need different angles to

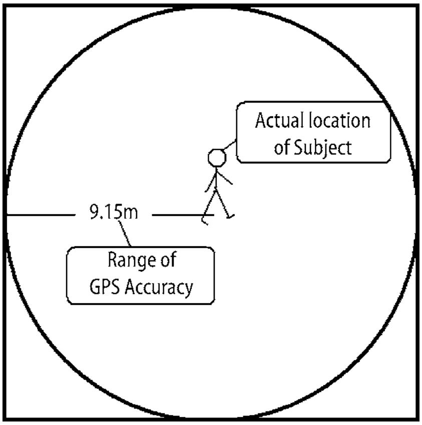


Figure 8.5. Range of GPS accuracy illustrated by man in bubble

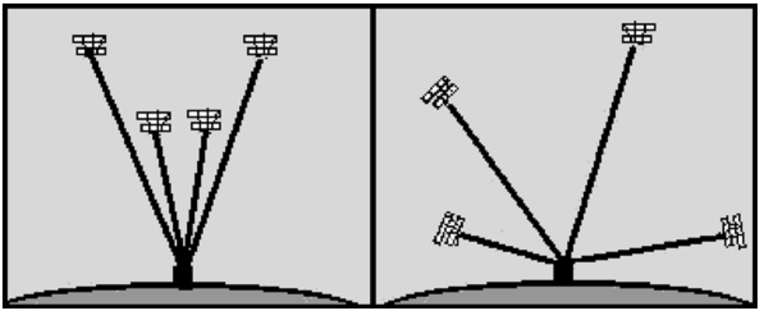


Figure 8.6. Satellite positions for accuracy: poor satellite positioning resulting in poor accuracy (*left*); good satellite positioning resulting in good accuracy (*right*)

decipher the precise location of the receiver (Figure 8.6). So, the receiver might be out in an open valley, but if all the satellites are right above it, the accuracy will diminish (Dana 2001).

Another factor that could influence GPS accuracy is the need for national security. The Department of Defense (DOD) developed GPS for military purposes. Since this technology is available worldwide, anyone can use GPS for location information. Originally, the government implemented a system called selective availability (SA) to scramble the signals from the satellites. This caused inaccuracies of 100 m. From May 2, 2000, U.S. President Clinton turned off selective availability but this decision could be reversed at any time (NGS 2000). However, this seems unlikely as the proliferation of civilian applications of the GPS permeate our society. Two new satellite systems are being implemented to augment the accuracy of GPS. The so-called Wide Area Augmentation System (WAAS) and the system known as EGNOS are geosynchronous satellites, which work in conjunction with ground stations to correct for some of the 9.15 m average error in the GPS system. These systems offer up to 3 m accuracy when the signals can be intercepted. Unlike GPS, WAAS cannot always be intercepted; even when one is in the open and at the time of writing EGNOS is not yet operational. Both systems are expected to increase coverage as more stations are installed.

Additionally, a European satellite version of GPS called Galileo is also expected to be operational by 2009. This all adds up to better accuracy, availability and reliability of global navigation systems. A new term in fact is emerging to refer to these multiple systems, global navigation satellite systems or GNSS.

8.2.3 Accuracy of GPS: Some Technical Issues

Several different techniques are used to try to enhance the accuracy of GPS devices. Three such techniques are discussed here.

Differential GPS

By placing a stationary GPS receiver at a known location, it can be used to determine exactly which errors the satellite data contains. This receiver acts like a static reference point and it is called a "Pseudolite". The stationary receiver can transmit an error-correcting message to any other GPS receivers that are in the local area. These additional receivers can use the error message to correct their positional solutions (Figure 8.7). This concept works because the satellites are so far above the earth that errors measured by one receiver will be almost exactly the same for any other receiver in a given area. This will reduce ionospheric, timing, and atmospheric errors, as well as any other errors introduced intentionally as was the case when selective availability (SA) was on.

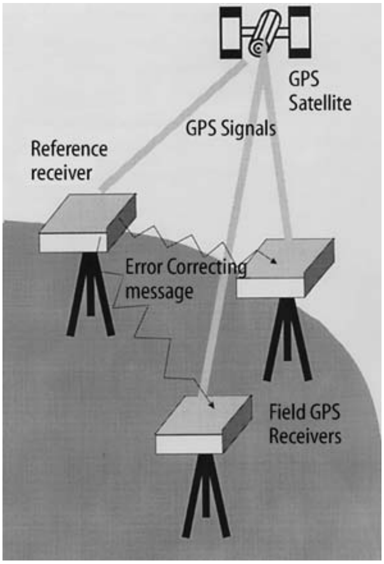


Figure 8.7. Differential correction: the Pseudolite principle

Before S/A was turned off this form of correction was becoming very popular and a number of companies began putting up the infrastructure needed to place these Pseudolites all over the country. But once S/A was turned off, subscribers of this service did not feel the improved accuracy offered was worth the cost of subscription, and as a result these companies either went out of business or specialized in specific applications, such as setting up a Pseudolite receiver to correct GPS errors in farming or surveying applications.

How trilateration can lead to an exact position

In GPS, when the distance to an object is known, that object's location can be represented by the equation of a circle. Without any other point of reference the location is merely somewhere on that circumference of the circle, at a radial distance away from its centre. The mathematical equation of a circle is

$$r^2 = (x - x_c)^2 + (y - y_c)^2$$

where r is the radius of the circle, distances x and y are the coordinates along the circumference of the circle, and x_c and y_c are the coordinates of the centre of the circle.

Now for a 2D solution, the distance to three other objects are required where the exact locations must be known. This will give three equations with two unknowns, and these can be solved for x and y to obtain the exact location of the object's position The three quadratic equations would given as

$$\begin{aligned}
(x - longitude_1)^2 + (y - latitude_1)^2 &= (distance_1)^2 \\
(x - longitude_2)^2 + (y - latitude_2)^2 &= (distance_2)^2 \\
(x - longitude_3)^2 + (y - latitude_3)^2 &= (distance_3)^2
\end{aligned}$$

Now for a 3D solution where altitude can also be determined, the location of a fourth object, specified through its radial distance, is required and then four equations can be solved for the three unknowns, latitude, longitude, and altitude.

The Kalman Filter

The Kalman filter is a multi-input/output recursive linear filter. It takes inputs from multiple sources and can estimate the state of the system based on its noisy outputs. The use of a Kalman filter system can reduce the error of the noisy outputs by real-time analysis of the mean-square of the estimated error.

The Kalman filter can also be used for combining multiple sources of location-based information and reducing its error by weighting each source accordingly as to reduce the error. The more inputs there are to the system the better the accuracy that can be obtained. The Kalman filter could be used to combine the inputs from GPS, DGPS, Inertial guidance, and WAAS to provide the best accuracy of a desired location.

8.2.4 Frequency Spectrum of GPS, Present and Future

The GPS system operates in the UHF (ultra high frequency) spectrum, which is between 300 MHz and 3 GHz, and within the L-Band frequency range. The L-Band consists of frequencies between (390 MHz–1.55 GHz) and is used for GPS satellites, satellite phones, SETI outer space exploration, and miscellaneous communication satellites. The primary frequency for civilian use is L1 = 1575.42 MHz. L1 carries a Pseudo Random Noise Course-Acquisition (PRN C/A) code, as well as an encrypted precision P(Y) code. L2 = 1227.6 MHz carries only the P(Y) code (Figure 8.8). The P(Y) code is used in military applications and the encryption keys needed to use the P(Y) code are rotated on a daily basis by the US government.

Since civilian applications only have access to the C/A code signal on the single frequency at L1, the accuracy is limited to approximately 10 m (without selective availability). Consequently, these applications cannot use dual-frequency corrections to remove the delay caused by the signals travelling through the ionosphere.

In December 2005, the first of the next generation GPS satellites was launched. The (Block IIR-M) satellite provides three new signals, two military signals (L1M, L2M), and a second civilian signal, L2C. Since the second signal for civilian use is in L2, civilian receivers can finally remove the ionosphere delay by using dual-frequency correction. In an open sky environment, this would probably improve the accuracy to less than 5 m.

Planned for launch in 2007, will be the second generation GPS satellite (Block IIF), which will add an additional frequency L5 = 1176.45 MHz.

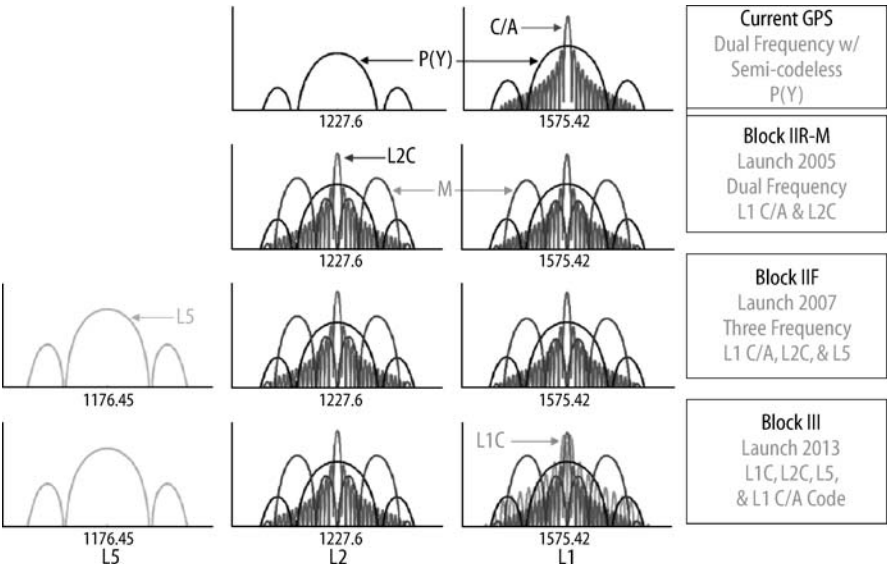


Figure 8.8. GPS signals: present and future

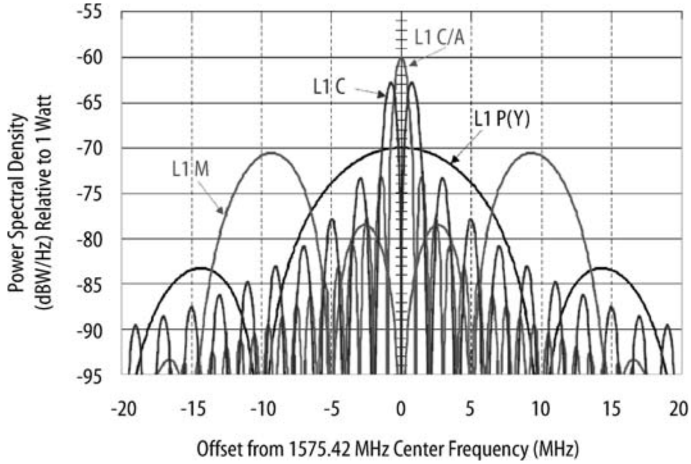


Figure 8.9. GPS signals at L1

Planned for launch in 2013 will be the third generation GPS satellite (Block III) that will add another civilian signal L1C. See Figure 8.9 for the proposed GPS signals at L1.

8.2.5 Other GPS Systems

8.2.5.1 Galileo System

Like GPS this is a global navigation satellite system constructed by the European Union. Galileo's first satellite GIOVE-A, was launched in December, 2005. Galileo is now scheduled to have a test constellation of four satellites by 2008, but no date for initial operating capability has been set, although it is assumed the system will be in operation by 2010.

Galileo is interoperable with the U.S. GPS system; it consists of 30 satellites in three Medium Earth Orbit planes at an altitude of 23,616 km. Each plane contains nine satellites plus one active spare. The higher orbital plane of Galileo results in better coverage over the northern parts of Europe where GPS coverage is weak due to its lower orbital plane. One revolution around the earth takes a Galileo satellite 14 h 4 min.

Galileo uses the same L1 frequency as GPS and has two signals L1F and L1P (Figure 8.10).

8.2.5.2 GLONASS – Russia

GLONASS (global navigation satellite system) consists of 21 satellites with three active spares in three circular orbital planes at an altitude of 19,100 km. Each satellite completes an orbit in approximately 11 h 15 min.

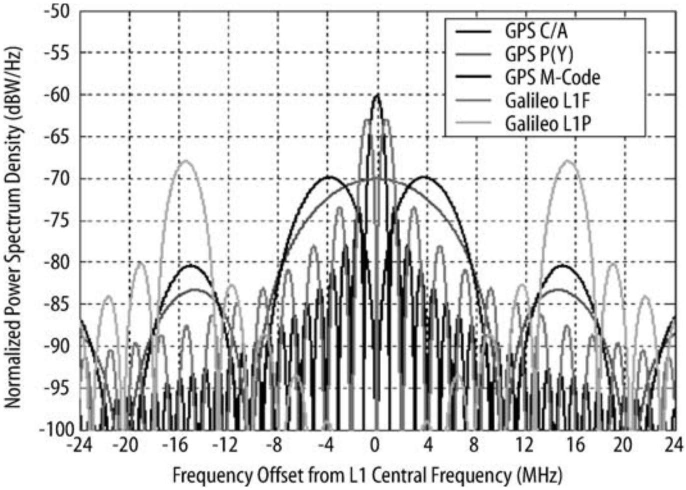


Figure 8.10. GPS and Galileo at L1

GLONASS was in full operation by 1995, but with the fall of the Soviet Union only eight satellites were operational by 2002. Three new improved GLONASS-M satellites, with a life span of seven years, were launched in 2004 with three more launched in 2005. The next generation of GLONASS-K satellites, which are lighter in weight, with improved functionality and a longer 10-year life span are to begin being launched starting 2006. Russia has said it plans to have the GLONASS operational by 2008 with 18 satellites covering all of Russia. A complete constellation of 24 satellites with full global coverage is planned by 2010.

Unique to GLONASS is that each satellite transmits its data at a different frequency determined by the frequency channel number of the satellite, thus allowing the user's receiver to identify the satellite. Two different carrier frequencies are used to transmit a precision (SP) signal and high precision (HP) signal, along with a standard C/A positioning code:

$$\mathrm{L1}: f_1(k) = 1602\,\mathrm{MHz} + k \times 9/16\,\mathrm{MHz}$$
$$\mathrm{L2}: f_2(k) = 1246\,\mathrm{MHz} + k \times 7/16\,\mathrm{MHz}$$

(where k is the channel number).

8.2.5.3 QZSS – Japan

Quasi Zenith Satellite System (QZSS) or Jun-Ten-Cho in Japanese is a modified geosynchronous orbit covering Japan and Southern Asia (Figure 8.11). The QZSS constellation will consist of three satellites moving in periodical highly elliptical orbits (HEOs) over the Asia region and should be functional by 2008. This system uses WAAS-like data on L1, L2, and L5. The satellites will have L, S, and Ku-band capabilities (L-band for positioning, S-band for broadcasts and low-speed communications, Ku-band for high-speed communications).

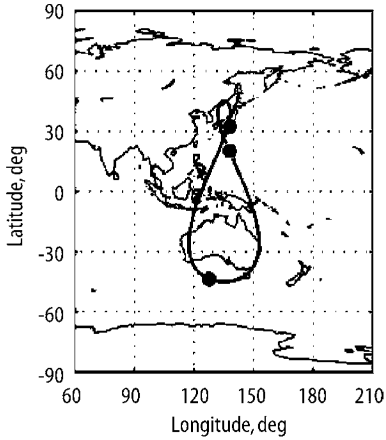


Figure 8.11. QZSS coverage

8.3 Application of GPS Principles

In the above section, the basics of how GPS works were presented. So, how do complicated numbers and distance calculations turn into helpful navigational information?

In GPS terminology, the latitude, 40° 44′ 52.4″ North, and longitude, 73° 59′ 6.12″ West, means that this point is currently 20,550 km (12,769 miles) away from satellite A, 19,312 km (12,000 miles) from satellite B, and 19,875 km (12,350 miles) from satellite C. Does that really help one navigate? Does one have a better idea of where they are going with that information alone? Probably not; however, with a GPS application and associated geographic databases, that particular latitude and longitude means that point is the Empire State Building in New York. Now that there is a name attached to those coordinates, you know exactly where that point is located.

Over the past few years, the GPS commercial market has grown rapidly. People are using GPS in rental cars, on hiking trips, and in many other recreational activities. Companies have been formed whose business it is to put names on latitude and longitude coordinates. These companies have constructed extensive databases with street names, addresses, business names, points of interest, restaurants, underwater wrecks, and the list goes on. Anything which is stationary is likely to be labelled. All of this extensive data has been put into electronic form so that it can be converted into various computer formats.

With this growth in electronic data, blind people no longer need be limited to the small amounts of location information they could previously glean from sighted people. The maps and points of interest are no longer just a drawing on an inaccessible print map. This electronic data can be converted into programs that work on computers designed for blind or visually impaired individuals.

The blind traveller can now be a co-pilot in a car, not just a passive passenger. They can keep the taxi driver honest; they can enjoy hearing about the sites and businesses passing by while in a car, bus, or even on foot. There is nothing more empowering for a blind person than getting around effectively and location information makes this much more possible for a lot more people.

8.4 Design Issues

Once the GPS and databases were firmly established as emerging mass-market technologies then three key aspects needed to be addressed for a truly accessible system for blind users: accuracy, portability and cost. These are all relative terms and there was considerable variation as to what level of accuracy was useful and what constituted portable; however, cost was more clearly defined. In the U.S., a product would have to be supported by government funding, which pays for the majority of equipment for people who are blind.

One of the first developments of a GPS product for blind and visually impaired people was initiated by Charles LaPierre as the part fulfilment of an undergraduate program at Carleton University; a fourth-year project entitled "Navigational System for the Visually Impaired", 1994 (LaPierre, 1993, 1998). From there, the project grew gaining support from two adaptive technology companies, VisuAide of Montreal, and Arkenstone of California. The project was informally called "*Sextant*" based upon the marine navigational aid for sailors, and a feasibility study was performed to determine the best hardware solution for the product as well as any potential pitfalls.

The design of *Sextant* began at Arkenstone, and two potential products were undertaken, "*Atlas Speaks*" the talking map for a personal computer and "*Strider*" the mobile talking map with GPS. Strider consisted of a laptop computer, GPS receiver and antenna, a speaker, and an external numeric keypad. At this time the package needed to be carried in a backpack (Figure 8.12). The GPS industry

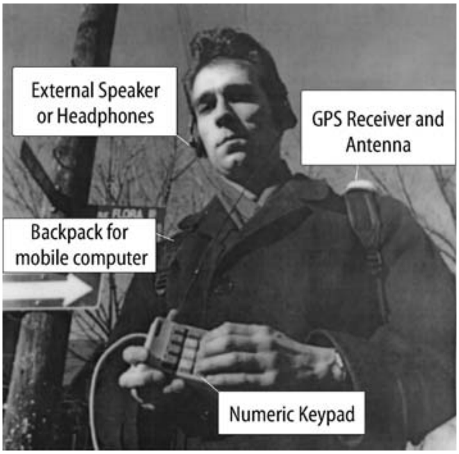


Figure 8.12. The Strider system

itself was still immature and the laptop was still too large and costly. The lack of integrated components made the systems bulky and cumbersome. Strider was never released to the public and its development was put on hold in 1997, around the time a similar project called MOBIC, being pursued in Europe, was discontinued.

Others too saw the utility of using GPS to assist in orientation. Jack Loomis proposed a research project on GPS-based navigation for the visually impaired in 1985 (Loomis 1985) and, with his colleagues Reginald Golledge and Roberta Klatzky, began developing the personal guidance system at the University of California, Santa Barbara in the early 1990s (Golledge *et al.* 1991; Loomis *et al.* 1994). This research continues to this day. VisuAide tested a GPS waypoint system called Mira, based upon their Magnum digital voice recorder. A number of individuals and companies in Europe and Japan were also working on projects but most prototypes remained in the university laboratory as size and accuracy discouraged most from commercialising a product.

In 2000, Mike May and Charles LaPierre, along with several other blind and visually impaired people, founded Sendero Group to resume development on the product formerly called Strider and to sell the Atlas talking map program for the personal computer. GPS companies were integrating the receiver and the antenna in the same device, roughly the size of a cell phone. Similarly, the larger computer companies were starting to produce smaller and more affordable laptop solutions. This made it possible for Sendero Group to release an updated version of Strider, called GPS-Talk, and the first commercial GPS product was on the market (Figure 8.13).

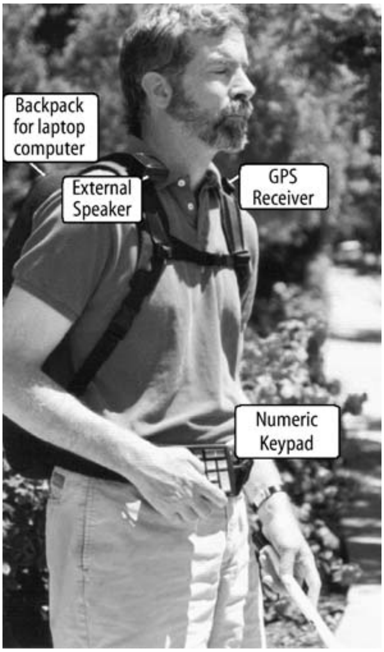


Figure 8.13. The GPS-Talk system

In the next couple of years, GPS-Talk amassed several hundred blind and visually impaired users. However, the consistent feedback from the consumers was a desire for an even smaller, more affordable system. Therefore, even though the size of the hardware had decreased substantially, this was still a cumbersome package to carry. Other drawbacks to the system were due mostly in part to the laptop limitations. Laptops only had at best 2–3 h of battery life, the boot up time for the system was a couple of minutes, and the computer might crash requiring the user to reboot. It seemed that GPS for blind individuals had hit a technological roadblock.

However, in the assistive technology industry, a new trend in portable note-taking devices was just starting. Various companies were striving to provide a device that would do basic word processing, scheduling, organizing contacts, emailing, and accessing the Internet. These devices are not unlike the Palm Pilot devices that were available to the sighted consumer market. Such devices are generically known as personal data assistants (PDAs). These new portable devices offered a platform for an accessible GPS application.

One of the only accessible portable devices, powerful enough to run GPS applications, was the BrailleNote and VoiceNote from Pulse Data International. In 2002, BrailleNote GPS was first released without street maps. This limited the information to points of interest (for example, restaurants, hotels, and museums) and routes that the user had input manually. In 2003, BrailleNote GPS Version 2 was released with street maps (Figure 8.14). This provided access to automatic route calculation and a full range of accessible GPS-based location information.

Other accessible PDAs would subsequently incorporate accessible GPS.

Alternate approaches called Trekker from VisuAide and StreetTalk™ from Freedom Scientific, also arrived in 2003 and 2005 respectively.

The Trekker is a stand-alone GPS module with dedicated PDA (WWW1 2006). As with BrailleNote/VoiceNote GPS, Trekker users can identify where they are, learn about area attractions, and find out how to get to specific destinations. A user can also record personal points verbally instead of having to type on the keyboard. However, the Trekker has only roughly 50% of the functionality of the BrailleNote/VoiceNote GPS. It offers an abbreviated version of accessible GPS to

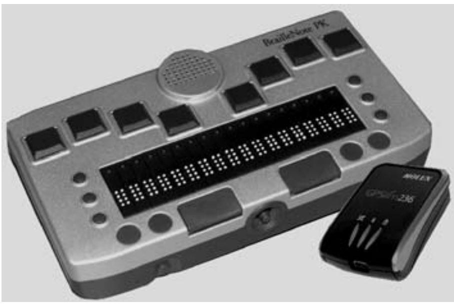


Figure 8.14. Photograph of BrailleNote GPS

a person who does not need or want an accessible PDA. The Trekker is very portable; it measures 87 mm × 134 mm × 27 mm (W'H'D) and weighs 600 g.

The StreetTalk™ is a GPS software addition to the PacMate PDA from Freedom Scientific (WWW2 2006).

Overview of accessible GPS systems

In this chapter, various accessible GPS systems from the past and present have been discussed. The accessible GPS industry is one of ongoing development, as the mainstream GPS technology improves, so will the accessible GPS solutions. Tables 8.1 and 8.2 reflect the GPS solutions available at the end of year 2004. These tables are not intended as a comparison of the various systems, but rather an overview of each system discussed in this chapter. Table 8.1 provides a commercial and technological overview, whilst Table 8.2 profiles the strengths and weaknesses of the various systems.

For a long time, navigation technologies were too complex, too large, and too expensive for the mass market. The advent of the accessible PDA has addressed the portability component of the design. GPS technology is now easier to develop and adapt, to make portable, and the price per unit has dropped dramatically because of the demand in the mass consumer markets.

Currently there is a mobile telephone GPS navigation system available in Europe called the Wayfinder EuroNavigator™. This system consists of a mobile telephone that is loaded with European maps and a Bluetooth GPS receiver (WWW3 2006). The user would load the appropriate map on their mobile telephone and the program would be able to give them directions and inform them of points of interest that are nearby. The issue of portability is definitely being addressed with mapping software adapted for a GPS enabled mobile telephone.

Table 8.1. Overview of accessible GPS systems: background to accessible GPS systems

Accessible GPS device	Strider	GPS-Talk	Personal guidance system (PGS)	BrailleNote GPS	Trekker	StreetTalk
Year introduced	1994	2000	1993	2002	2003	2005
Currently available to public?	No	No	No	Yes	Yes	Yes
Released to public	No	Yes	No	Yes	Yes	Yes
Developing agent	Arkenstone	Sendero Group	Loomis *et al.*	Sendero Group & Pulse Data International	Visuaide	Freedom Scientific
Computer- or PDA-based	Computer	Computer	Computer	PDA	PDA	PDA

Table 8.2. Overview of accessible GPS systems: strengths and weaknesses of accessible GPS systems

Accessible GPS System	Computer- or PDA-based	System strengths	System weaknesses
Strider	Computer	Identical in functionality to GPS-Talk, but was never released to the public	GPS receiver and portable computer size not miniaturized enough to make a marketable product
GPS-Talk	Computer	First accessible GPS system released to the public; comprehensive orientation software; worked on any Windows-based computer including miniaturized wearable computers; self-voicing application; no map loading limitations	Size, weight, and cost prohibitive; portable computer limitations such as poor battery life, software crashes requiring computer rebooting, time to access information could exceed 3 min; route calculations limited to 5 miles; no concern for one-way streets and highway travelling
Personal guidance system (PGS)	Computer	Includes electronic compass providing additional location information; differential GPS allowing for better accuracy; headphone and microphone voice recognition system	Not released to the public; size and weight prohibitive – separate GPS receiver and antenna; not intended to be a compact package; not safe for independent pedestrian travel, as headphones obstruct hearing environmental noises such as cars
BrailleNote GPS	PDA	Improved upon the powerful software features found in GPS-Talk; no route length limitation; full access to the maps with or without GPS; fully accessible PDA with Internet access, email capabilities, word processing, address book, mp3 player, *etc.*	Compatible only with specific Windows CE device – the BrailleNote or VoiceNote
Trekker	PDA	Stand-alone GPS unit; small package and light weight; provides street, point of interest, and directional information; uses voice recording to name points of interest	Routes limited to one map section
StreetTalk	PDA	Works with the PacMate PDA; is an inexpensive software addition to the PacMate	Software is adapted from a GPS program for sighted people and not all functions are accessible; does not allow for input of personalized points of interest or routes

Although development funding has improved, it remains a significant issue; positional accuracy has been addressed with WAAS and EGNOS and will continue to improve with other GPS augmentation systems. Portability has improved significantly since the days of Strider in a backpack and units will continue to get smaller.

8.5 Development Issues

Development of a GPS system for a blind or visually impaired individual requires certain transformations to the mass-market technology designed for sighted people. We cannot just take a GPS system out of a car, even those that talk, and hope to use it for navigation by blind people.

The main issues to be explored are:

- Choosing an appropriate platform for the device.
- Choosing the best GPS receiver.
- Creating a package that makes the system easy to carry.
- Deciding the issue of integration *vs* stand-alone.

8.5.1 Choosing an Appropriate Platform

In an ideal world, the GPS systems created for the sighted market would have a verbose setting for a blind user so that all the screen information could be accessible. Some of the commercial car GPS systems have speech output but the menus and settings displayed on the screen are not announced or described in speech.

In general, when a blind person uses a computer device, there are various ways to input the information, voice input, computer keyboard input (typing), or Braille keyboard input. Although voice recognition technology works reasonably well in indoor situations, it is less reliable in a noisy outdoor environment and the addition of a microphone adds complexity to the system. So that leaves either typing using a computer or a Braille keyboard. The choice between the two different types of keyboards is a personal one, for every person has preferences. It is advantageous to have the flexibility to provide both solutions.

There are two options for delivering the information to blind users, speech and Braille output. There is a conundrum with regard to speech output. Earphones are fine while listening in a vehicle or while stationary; however, a person's ears should not be covered when walking or important mobility information could be obscured. Hence, a loud speaker is the practical trade-off. This speaker is best built into the PDA so as to avoid extra weight and wires, amplification and battery power.

Future interface solutions may include vibrating tactile or, more likely, Bluetooth earphones or sound tubes near, but not obscuring, the ears. Perhaps the user would wish to use wireless headphones while sitting on a bus bench. Since 3D location information has been shown to be useful, especially for displaying complex surroundings, there is reason to explore further the different ways to deliver this dimension to users.

For those who read Braille, it is helpful to have the information in refreshable Braille. Braille displays are designed to output one line of text at a time. When the user is finished with the line that they are reading, they can advance to the next line. This is analogous to scrolling text on news broadcasts. Unfamiliar and complex street names are easier to feel in Braille than listening to speech.

Many people use speech and Braille together. There are advantages to using speech, Braille, or both. In general, speech is the faster way to receive bulk information and Braille is better for reading details. Those accessible PDAs, which provide both, are best. Refreshable Braille does add weight and two to three times the cost of speech to a system.

8.5.2 Choosing the GPS Receiver

Since the introduction of the first consumer GPS receivers in 1995, the metamorphosis of the GPS receiver has been dramatic. The variety of GPS receivers available includes receivers with different connectors (Serial, USB, Compact Flash, *etc.*), receivers of different sizes (from TV remote size to matchbook size), many different user applications (screens, maps, points of interest files), and so on. So which features are important when constructing a GPS system for use by visually impaired and blind individuals?

A small sized, lightweight device is important for the accessible GPS package. Durability, price and functionality must be considered when evaluating the tradeoffs with size.

Another important issue with GPS receivers is the accuracy reliability. A receiver must work well when there is marginal satellite visibility, for example, in the window of a bus or on the sidewalk close to a tall building. As stated before, the GPS information from the satellites has its limitations, but there are receivers that more efficiently capitalize on the information provided by the satellites. It is important to test various receivers and judge which ones are more reliable in marginal situations.

Many receivers offer a plethora of functions for sighted users. Unless they are made accessible, screens and programming buttons are of no advantage to a blind person. The optimal GPS receiver for this GPS system needs to be easy to use. It should have a distinct ON/OFF button, programming that requires little to no input from the user and an audible alert for low battery.

Durability is very important as blind users use the GPS more than the average recreational sighted user. Blind people are constantly putting the unit on and taking it off, using it in a vehicle and while walking. Cables and connectors should be rugged and eliminated when a wireless option is available.

8.5.3 Creating a Packaged System

Once a GPS receiver is selected, it is important to remember that this will be a system that is carried by its users and not left stationary in a vehicle like most systems for sighted users. Consider a blind person travelling down the street. The first thing that you might associate with that pedestrian is a long cane or guide dog. Both of these mobility aids require the use of a hand, leaving the other hand free

Figure 8.15. Packaging the GPS system. Note: one hand is free to operate the GPS program

for operating the GPS system. One configuration is where the GPS system is carried like a shoulder bag and the GPS receiver clips into a shoulder strap (Figure 8.15).

Another benefit of having the system configured as seen in Figure 8.15 is that it can be efficiently removed and stowed when not in use. Once a visually impaired or blind person has reached their destinations, they probably do not want to continue to carry a GPS system.

8.5.4 Integration *vs* Stand-alone

As with most product acquisitions, there are cost benefit trade-offs when choosing an accessible GPS solution. When GPS is combined in one unit with other applications, the individual application costs go down but the integrated unit it is part of is probably more expensive than a stand-alone GPS unit. The same is true of the smaller incremental size of the GPS when part of an integrated unit. There are so many accessible devices that a blind person is forced to choose how many and which units to carry. These devices might include a cassette recorder for books on tape, a memo recorder for quick voice messages, a mobile phone, a PDA or laptop, a compass, a magnifier for a low vision person, maybe an obstacle detection device like the Miniguide and of course, GPS. It is very useful when any of the functions of these devices can be combined in a single unit because it reduces the incremental cost and size of each application. A similar situation exists in the mainstream GPS market there are dedicated GPS units and PDAs with add-on GPS applications to choose from.

Functionality does not conceptually have to correlate with stand-alone or integrated units. Historically, if you wanted a top of the line sound system, you purchased separate components. This does not apply to computer-based technology. Functionality is more likely to be linked to the memory capacity and processing speed of the computer chip running the GPS. Functionality will have the most to do with the maturity and sophistication of the software. Sendero's GPS software, for example, has been in the marketplace in one form or another since 1995 giving it the benefit of user input over many years. Newer GPS applications primarily developed for sighted users have basic functionality but are lacking features specific to blind users.

Mobile phones with built-in or Bluetooth GPS are likely to offer the convenience and cost benefit of integration with basic functionality. They have the benefit of more up-to-date map data residing on a server. However, the mobile phone GPS is not likely to have the power of the accessible PDA GPS for narrating location information and providing "look-around" functionality.

8.6 User Interface Design Issues

Designing a device that provides location information requires careful research and planning for the user interface. The main objective is to deliver relevant orientation information while not overloading the user. Since the definition of what is considered *relevant* to the end-user is different for every individual and travel situation, the user interface must be configurable but not complicated. The balance between configurability and complexity must be driven by user input. The more experience a developer has with user feedback, the better this balance will be achieved.

8.6.1 How to Present the Information

Everyone has a different preference when it comes to obtaining travel directions; this is also true for blind travellers. Generally, a blind person will get directions in two different forms. The first form uses a sound that is emitted by the source of the object. For example, to find a person in a room, a blind person might follow the sound of that person's voice. This type of auditory cue will be referred to as *3D Sound*. The other form of directional information is the type of directions obtained when on the street asking for directions. For example, "Big Mama's Pizza is ahead of you 8 m away." Other variations of this type of information use clock face references. So the same example from above would be "Big Mama's Pizza is at 12 o'clock, 8 m away." These auditory cues are referred to as *Spatial Language*, because they use language to describe where an object is located. According to research studies from the University of California, Santa Barbara, 3D sound and spatial language are roughly comparable in their effectiveness as auditory interfaces. For guiding a person along a route, 3D sound delivered by a virtual acoustic display led to somewhat better performance and user evaluations than spatial language (Loomis *et al.* 2001). With regard to monitoring and reporting off route locations, 3D sound and spatial language are very similar in their effectiveness (Loomis *et al.* 2002; Klatzky *et al.* 2003) but building up mental representations of these off-route locations takes a little longer with spatial language. Although 3D sound is slightly more effective as an interface, it suffers from these disadvantages when implemented using a virtual acoustic display: earphones are necessary, and specialized hardware is needed to implement the display.

8.6.2 When to Present the Information

Now that the best way to describe the directions has been determined, the best time to receive directional information will be considered. Again, personal preference plays a role in this issue. Some people like to receive all their directions at once and others like to be given announcements as they travel the route. Some blind people have trained to use their memory extensively when it comes to directions; however, it is still nice to have reassurance while on a route. The GPS program should have the flexibility to be able to provide information to suit the user and the situation. For example, the announcements can be quiet until the traveller approaches the destination or when a turning point or significant point is approached. On the other hand, announcements can be set to repeat every 10 s, giving the user a continuous commentary on the upcoming intersections or remaining distance to the destination. Automatic announcements keep the user from having to find a key or push a button, especially if that hand is occupied carrying something, for example a mobility device like a long cane. Manual functionality is important for the user who does not want his or her thoughts interrupted or who is concentrating on mobility information.

8.6.3 What Information to Present

Finally, it is important to identify the type of information that a blind traveller will need. Using GPS technology, every point on earth becomes an electronic landmark; every point of the Earth attains a unique name. So, there is no limit to the amount of information that can be announced. Filtering the important information for the given user in a given situation becomes the challenge. The essential location information like street names, upcoming intersections, businesses, cities, and landmarks should be the first priority in creating a GPS system designed for blind travellers. Using the example from the first section of the chapter, if the task was to find Big Mama's Pizza, it would help to know that Big Mama's pizza is at the intersection of Broad Street and 22nd Street. It would also be helpful to know that it is ahead of you and 38 m away. With a GPS system it is also possible to know your speed of travel, your heading direction, your distance travelled, and that you have not wandered in the wrong direction.

8.7 Test Procedures and Results

Because GPS technology is so new for blind and visually impaired users, the formal testing pursued relies mostly on a select group of advanced users.

The first stage is a design stage, which consists of an evaluation of functions and features recommended by users of earlier versions. These suggestions and the frequency of each suggestion is monitored and archived. Feedback is recorded from blind users and from teachers and orientation and mobility instructors. By the time work begins on the next version of the program, a good idea of the priorities and concerns of existing users has been obtained and analyzed.

Based on this archive of experience, "stories" which reduce ideas and concepts to real user situations are constructed. A team of users, engineers and consultants from the assistive technology field evaluate these stories. Once the stories are prioritized, they can then be entered into a bug and feature-tracking database, which allows engineering and product management to closely track development, to reprioritize and work through the software design.

The alpha stage of testing is done by a core group of blind users who are technically competent. In the alpha stage, features are added and modified based upon feedback from the testers. The software is not assumed to be fully functional during the alpha stage. The goal is to implement most of the features and to fix bugs from previous versions during the alpha stage.

The beta testing cycle is designed to test the software on a variety of users, both technical and novice users from a geographical cross section from the U.S. and other countries if appropriate. We wish to see how the software performs under a wide range of usage for at least two months.

During the beta cycle of tests, the purpose is to finalize features and commands based upon feedback from users. We do not normally change fundamental aspects of the program during beta testing.

Once a stable beta version is available, then moves are begun to release the program to the public.

8.8 Case Study in Commercialisation

When developing a new technology to enhance the orientation of blind and visually impaired individuals, there are various factors that need to be addressed:

- The consumer must understand the value of the technology.
- The limitations of the technology must be stated clearly.
- Developing the technology is an ongoing process.

8.8.1 Understanding the Value of the Technology

Assimilating copious location information may be both exciting and challenging at first for a blind user who has traditionally only had limited street map information.

When an intersection is described in terms of the clock face or right-left-front-back, the user may not initially have the mental imaging skills to understand the intersection description. It takes practice hearing these descriptions before the user can automatically assimilate the information. Users improve their cognitive imaging skills after becoming proficient with the GPS and digital maps.

Although accessible GPS has begun to be used, there is some lack of awareness and understanding of the empowering value of location information. If you have never had something, how can you appreciate its value?

Some people feel that there is no need for advanced technological devices that tell you where you are in space; listening and walking with a long cane or guide dog are sufficient. However, way-finding orientation technology is not designed to replace the long cane or the guide dog. Wayfinding orientation technology provides a blind traveller with more options. These electronic tools will be used primarily to gather detailed and accurate information about the outdoor environment. With these new technologies, the blind traveller can become safer, more efficient and more independent especially in the complex modern world.

8.8.2 Limitations of the Technology

It is fair to say that the current GPS technology offers users many benefits but there are also limitations, issues like accuracy and seamless availability.

On average, commercial GPS receivers are accurate within 9.15 m (30 feet). This accuracy will not guide a blind traveller straight in the front door of a restaurant without exploration by the user when near the destination. Instead of being a system that pinpoints the location, it is a system that gives access to location information that was previously not available to people with visual impairments. It creates a general description of the surroundings and helps in the *cognitive mapping* process that is necessary for travelling successfully and independently. Cognitive mapping is a process in which an individual acquires, stores, recalls, and decodes information about the surrounding environment (Downs and Stea 1973). It is best to start integrating location information training into the lives of blind people rather than waiting for perfect worldwide GPS with centimetre accuracy.

One of the greatest weaknesses for GPS-based navigation systems has been in areas where the GPS signals are not available, for example inside buildings and subways. A friend recently asked "When the boat was first introduced, what if people had said, but it doesn't travel on land?" What is now available on the BrailleNote GPS is very beneficial; it is the beginning of emerging GPS technologies for blind and visually impaired individuals. Those who wait for the perfect solution and do not start learning how to utilize location information now to further their orientation and understanding of the environment will be missing the boat.

8.8.3 Ongoing Development

As new technologies evolve, there is a critical point when that technology passes from fringe to mainstream. Key to development in a niche market, especially a small one like assistive technology for visually impaired and blind people, is adapting a mainstream technology and not a fringe one. Since cost is such a huge barrier for a relatively low-income population, piggybacking mainstream technology is fundamental to reducing the cost barrier and to ensuring that the adaptive technology will keep pace.

Once a GPS system was released for sale in the assistive technology market, improvements in the functionality, accuracy and convenience was then the focus.

It is important to cultivate communication between end-users and developers. Gathering user testimonies on how the technology did and did not work in certain situations drives the need for improvement in future versions.

As GPS receivers become wireless, smaller and more accurate, these commercial devices must be integrated into the adaptive technology. This is best done by keeping the GPS hardware separate from the adaptive technology as opposed to building it into a dedicated Braille or speech unit. There are always trade-offs, which influence the decision whether to go with separate components versus integration but it is ideal to minimize the adapted part of the solution, while maximizing the commercial components. This will be balanced by the priorities of functions blind users require.

Also, in the near future, augmented GPS systems with enhanced positional accuracy will become available. Coupled with more accurate street maps, we will be able to pinpoint more reliably the location of a target destination. There are both satellite and ground-based systems, which correct for the GPS errors. Sub-meter and even centimeter accuracy is currently available but the costs and size restricts it to commercial applications.

Indoor navigation is another important area for improvement. Sendero Group has been working to integrate an indoor navigation module with GPS to enhance navigation in urban canyons and indoor environments. Motion detectors in a personal navigation module (PNM) including a compass, keep track of where the user is walking. Waypoints and points of interest can be triggered by this indoor system in the same way that the GPS system does for outdoor environments.

Talking Signs™ using infrared transmitters are used to announce doorways and other key indoor locations through a hand-held audio receiver. An infrared beacon system has been developed in the Czech Republic and is being tested in some European cities. Talking Lights™ is another indoor technology under development. In this technology, a modified ballast can be inserted in a florescent light fixture to give it a unique position identification. These sensors can theoretically be connected to a PDA, which has a database of a building, to trigger indoor location information.

For this indoor technology to be effective databases of indoor information must be developed and occasional "known" positions must be part of the indoor environment. These known indoor locations can work in conjunction with the motion sensing system, PNM, to keep it accurate. When coupled with GPS, this has the potential for a seamless indoor/outdoor location information and navigation system.

Telephone and PDA manufacturers now have integrated GPS mobile phones. These GPS cell phones will not replace the need for navigation narration from a PDA, but they will provide the ability to spot check locations and to receive periodic verbal directions. For example, I press a speed dial number, which connects me to a map server. My GPS position is sent by my phone to the server. Now, I can press "1" for nearest important location points or "2" for the nearest street intersection. The same type of information currently presented on the PDA can now be accessed by the GPS cell phone and spoken back to the user by text-to-speech software.

8.9 Chapter Summary

The difficulty of travel for the visually impaired or blind person is a significant barrier for social inclusion and personal independence. In a simple decomposition, it comprises two main areas; one is obstacle avoidance and the other is long distance navigation. Considerable thought and development has been given to trying to provide assistive technology for the two problems. In this chapter, the use of GPS technology as a contribution to the solution of the navigation problem was described and discussed.

The chapter opened with a discussion of the context of the navigation problem where issues, such as the importance of location information and the difference between mobility and navigation were considered. This led into a section on the basics of global positioning system technology; its principles, its accuracy and its shortcomings were introduced and discussed.

The middle part of the chapter was devoted to the explaining how these GPS principles are used in navigational assistive technology aids for visually impaired and blind people. Consideration was given to the design issues, how the new device should be packaged, the all important user interface and finally, the test procedures and the outcomes. Particular attention was given to the interface design as the success and adoption often crucially depends on an interface that actually meets the needs of the target end-user community. Most of this material was based on the authors' personal experience in bringing such a navigation assistive device to the marketplace.

The final section of the chapter examined commercialisation issues and reported the authors' experience of marketing a navigation assistive technology device. This section concluded with a brief discussion of how ongoing technological developments are likely to affect future solutions of the navigation problem for visually impaired and blind people.

Questions

Q.1 What is location information and why is it important?
Q.2 What are the differences between mobility and orientation devices?
Q.3 What are the conditions that influence GPS accuracy?
Q.4 How can GPS be used to give location information?
Q.5 What are the primary strengths and weaknesses of GPS-based navigation?
Q.6 What factors are important to creating a user interface?
Q.7 Research the systems that augment the GPS accuracy, WAAS and EGNOS. How do these systems provide better accuracy?

Projects

P.1 If you are sighted, experience what it is like to travel without sight. Blindfold yourself and have a person that you trust lead you around a familiar place and an unfamiliar place. Your guide should not give you any information about your surroundings. Record your observations. How did travelling differ without the use of sight?
P.2 Now repeat the same exercise as above, but have your guide talk to you about the environment through which you are passing. How did this differ from the first experience? Were you more confident? Did you start to make a cognitive map of the area?
P.3 Use indoor navigation devices to map a building. Mark indoor points of interest and transitions in hallways or large rooms. Build indoor databases and explore which information is most useful.
P.4 There are a number of sensors and sources of location information that have not been discussed in this chapter, either because they are not commercially viable yet or because they have not yet been invented. Wi-Fi for example is a commercial technology that may be able to play a role in delivering dynamic location information or position. Explore this and other solutions to augment content and accuracy.

References

Dana, P.H., 2001, Global Positioning Systems Overview, The Geographer's Craft Project, Department of Geography, The University of Colorado at Boulder. http://www.colorado.edu/geography/gcraft/notes/gps/gps_f.html

Downs, R.M., and Stea, D. (Eds), 1973, Image and Environment: Cognitive Mapping and Spatial Behavior, Aldine, Chicago, U. S. A.

Golledge, R.G., Loomis, J.M., Klatzky, R.L., Flury, A., and Yang, X., 1991, Designing a personal guidance system to aid navigation without sight: progress on the GIS component, International Journal of Geographic Information Systems, Vol. 5, pp. 373–395

Klatzky, R.L., Lippa, Y., Loomis, J.M., and Golledge, R.G., 2003, Encoding, learning, and spatial updating of multiple object locations specified by 3-D sound, spatial language, and vision, Experimental Brain Research, Vol. 149, pp. 48–61. http://www.psych.ucsb.edu/~loomis/ or http://www.geog.ucsb.edu/pgs/main.htm

LaPierre, C.M., 1993, Navigational System for the Visually Impaired, Faculty of Graduate Studies and Research, Department of Electronics, Carleton University, Ottawa Canada., http://lapierre.jammys.net/nvi/index.html

LaPierre, C.M., 1998, Personal Navigation System for the Visually Impaired, Department of Electronics, Carleton University, Ottawa Canada. http://lapierre.jammys.net/masters/index.html

Loomis, J.M., 1985, Digital map and navigation system for the visually impaired, Unpublished paper, Department of Psychology, University of California, Santa Barbara, http://www.psych.ucsb.edu/~loomis/

Loomis, J.M, Klatzky, R.L, Golledge, R.G., Speigle, J.M., and Tietz, J., 1994, Personal Guidance System for the Visually Impaired, Proceedings of the First Annual ACM/SIGGAPH Conference on Assistive Technologies, Marina Del Ray, CA, October 31-November 1, 1994, pp. 85–91, New York: Association for Computing Machinery. http://www.psych.ucsb.edu/~loomis/

Loomis, J.M., Golledge R.G., and Klatzky R.L., 2001, GPS-based navigation systems for the blind, pp. 429-446 in Fundamentals of Wearable Computers and Augmented Reality, W. Barfield and T. Caudell (Eds), Mahway, NJ, Erlbaum http://www.psych.ucsb.edu/~loomis/

Loomis, J.M, Klatzky, R.L., and Golledge, R.G., 2002, Spatial Updating of Locations Specified by 3-D Sound and Spatial Language, Journal of Experimental Psychology: Learning, Memory and Cognition, Vol. 28, No. 2, pp. 335–345. http://www.psych.ucsb.edu/~loomis/

Millar, S., 1988, Models of sensory deprivation: the nature/nurture dichotomy and spatial representation in the blind, Journal of Behavioural Development, Vol. 11, pp. 69–87

NGS, 2000, Removal of GPS Selective Availability, National Geodetic Survey. http://www.ngs.noaa.gov/FGCS/info/sans_SA/

TNL, 2001, Trimble Navigation Limited, 2001. http://www.trimble.com

WWW1, 2006, Trekker, http://www.humanware.ca/web/en/p_DA_Trekker.asp

WWW2, 2006, StreetTalk, http://www.freedomscientific.com/fs_products/StreetTalk.asp

WWW3, 2006, Wayfinder EuroNavigator™, www.wayfinder.com

Additional reading

El-Rabbany, A., 2006, Introduction to GPS: The Global Positioning System, 2nd Edition, Artech House Publishers, London, UK

Kaplan, E., and Hegarty, C., 2006, Understanding GPS: Principles and Applications, 2nd Edition, Artech House Publishers, London, UK

Prasad, R., and Ruggieri, M., 2006, Applied Satellite Navigation using GPS, GALILEO; and Augmentation Systems, Artech House Publishers, London, UK

9 Electronic Travel Aids: An Assessment

Learning Objectives

Many assistive technology mobility products are expensive. Visually impaired people are a community group with considerable variability in their range of sight abilities. Thus, it is very useful to learn about the relative successes and limitations of many of the currently available mobility products; this chapter provides the framework for such an assessment.

The chapter opens with an analysis of the various types of methods that can be used to set-up an end-user assessment exercise. Part of this concerns the framework of user requirements of the products to be assessed and another part is concerned with selecting the way of collecting the raw data of end-users responses and experiences.

The second part of the chapter presents the author's findings for an end-user assessment of six obstacle avoidance mobility aids and two accessible navigation aids. The chapter closes with a discussion on the importance of training to achieve the best return from the use of advanced technology to assist in mobility and navigation.

Thus, the emphasis of this chapter is not on the technology but on user needs and the interactions between user, technology, task and environment. By the end of this chapter you should:

- Understand why assessments of electronic travel aids should be user-focused.
- Be able to identify a number of techniques that may be used in user-focused assessments and user requirements gathering.
- Be able to identify key advantages and disadvantages with these methods.
- Understand the distinction between mobility and navigation aids.
- Understand the distinction between primary and secondary travel aids.
- Be able to identify key user-focused advantages and disadvantages of a selection of electronic travel aids.
- Understand some of the user-focused strengths and weaknesses of modern-day electronic travel aids in general.

- Understand some of the key problems with the current lack of specialised training in the use of electronic travel aids.

9.1 Introduction

Research and development into electronic travel aids for blind and visually impaired people has been under way for around four decades. From the earliest days of this research to the present there have been many technological successes. Today we see refinements of some of the early electronic travel aids being used in modern devices and we see new innovative technologies being developed. Despite the technological successes, few blind or visually impaired people use an electronic travel aid.

Why is it that there are electronic travel aids out there, which perform well technologically, yet their take-up by potential users, that is blind or visually impaired people, is low? Do blind and visually impaired people not have difficulties travelling? Do they not want devices to assist their travel? Why is it that these technologically successful electronic travel aids are not being used by blind and visually impaired people?

Some blind and visually impaired people can and do travel safely and independently but many do not. According to Golledge (1993), independent travel and interacting with the wider world is the biggest hurdle that blind and visually impaired people face. A survey of the mobility of blind and partially sighted people in Nottingham found that three in ten blind or visually impaired people never go out alone (Clark-Carter *et al.* 1981). The number of older people with vision loss who never go out alone is even higher, at around two in every three elderly persons (Yerasimou 2002). As many as 86% of blind and visually impaired people feel that difficulties using public transport have adversely affected their job opportunities (Campion *et al.* 2003). Yerasimou (2002) reports that accidents, such as knocks, trips and crashes are commonplace for blind and visually impaired travellers and that the fear of such accidents and the fear of becoming disorientated or lost deter many from independent travel. There is, then, plenty of evidence to show that independent travel is a challenge for blind and visually impaired people. So, why are electronic travel aids not in great demand?

There are, it seems, many factors that contribute to the low take-up of electronic travel aids by blind and visually impaired people. Technological performance is only part of the equation. Technology does not operate in isolation, it must be considered within the broader context. Users interact with technology to perform tasks within a social, economic, political and physical environment. All these factors need to be considered when assessing technologies, such as electronic travel aids.

As an example, consider the scenario of a blind long cane user walking along a familiar platform at a busy train station. In this case, the user, that is the blind person, knows the way, the technology, that is the long cane, is well suited to detecting obstacles and hazards including the edges of platforms, and thereby facilitates the task of safely walking along the platform. What happens, though, when other passengers stand in the way? The long cane hits them like any other obstacle. Whilst most people will accept an apology and be understanding some may be angered by being whacked around the ankles by a long cane. What was a simple task of walking along then becomes an uncomfortable social interaction. Now think about what happens if a well-intentioned person wants to help our user but does not know how to go about helping? Taking a true anecdote, in one such instance the eager-to-help sighted person took hold of the tip end of the user's long cane and attempted to steer our user along. This rendered the long cane useless. No longer could it detect obstacles, this method of attempting to guide is ineffective, thus the task became impossible and changed to one of explaining to the sighted person that help was not required and persuading him, to let go of the long cane. Additionally, if a train were to arrive, creating substantial amounts of noise, this may interfere with the user's ability to listen to the echoes created by the sound of the cane tip, thus reducing the user's ability to use echolocation to hear the location of obstacles and gaps. This example shows how physical and social factors can influence users of travel devices. These factors, and others, must be taken into account by assessments of travel aids.

This chapter looks at user-focused assessment of electronic travel aids. First, it considers a number of methodologies for determining user requirements, of performing user-focused evaluations of electronic travel aids and considers some of the methodological issues that need to be addressed in the design of research involving participants with visual impairments. Second, it takes a look at how well some modern-day electronic travel aids meet the needs of potential users. Finally, it considers the requirements users have for training in the use of electronic travel aids. The emphasis of this chapter is not on the technology but on user needs and the interactions between user, technology, task and environment.

9.2 Why Do an Assessment?

Before embarking on any assessment of electronic travel aids it is crucial to answer the question: Why is this assessment being done? What is it purpose? An assessment may be intended to enable comparison between different electronic travel aids. It may be to establish how well suited, to the needs of users, a particular device is. It maybe to establish how much benefit users can experience from using the device. Perhaps, though, the most valuable assessments are those that seek to identify strengths and weaknesses of one or more devices. This information can be used to guide future developments, ensuring that the development process is streamlined to the target of meeting user needs. This kind of formative assessment, or profiling, records both positive and negative points. Subsequent development can act upon this information. Thus, this is an important element in designing electronic travel aids of the future. It helps to ensure that future devices will be both technologically successful and more widely used. This kind of assessment will usually need to be thorough and, therefore, in the short term may appear a costly exercise. The benefits of such comprehensive formative assessments, however, will be realised on later iterations of the design cycle.

The assessments presented in this chapter are intended merely to identify key advantages and disadvantages of a range of travel aids and consider some of the issues that may arise.

9.3 Methodologies for Assessments of Electronic Travel Aids

User-focused assessments of electronic travel aids can be performed in many different ways using a range of methodologies. Each approach has its own set of advantages and disadvantages. For each assessment, therefore, it is important to choose the method or combination of methods that are best suited to the purposes of the individual assessment. Only a brief introduction of some of the techniques that may be used for user-focused assessments is given here. It is beyond the scope of this chapter to discuss these methodologies in detail but more information on the methods presented here and on other related methods can be found in the books by Oppenheim (2000) or Wilson and Corlett (1995). Let us consider first methodologies for eliciting user requirements and the use of these in heuristic, or expert evaluations, and then methodologies for hands-on user assessments.

9.3.1 Eliciting User Requirements

Several methodologies exist for eliciting user requirements. These include questionnaires, interviews, focus groups and protocol analysis amongst others.

9.3.1.1 Questionnaires

Questionnaires involve a structured set of questions. They may be administered in writing or as a structured interview. Questionnaires have the advantage of allowing large amounts of data to be collected whilst demanding relatively few resources. Primarily for this reason they can be a useful tool in gathering user requirements. However, in the context of eliciting user requirements for electronic travel aids they have a number of disadvantages. One important drawback, is that such a static method is not well suited to enabling people to think about their needs in a travel setting. This is especially significant when we recall that few blind or visually impaired people have experience of using any electronic travel aid. Indeed, many rarely if ever travel alone. This should not be underestimated and should be borne in mind later in this chapter when we consider some currently available electronic travel aids. This is further confounded by the lack of interaction between researcher and respondent. Questionnaires do not allow for discussion between researcher and respondent and therefore do not enable the researcher to clarify concepts or pursue a topic further. Questionnaires, then, are a useful means of gathering large amounts of data but may not provide all the data that is needed. They may be a valuable first step but for most purposes other techniques will be needed to augment and validate data obtained through questionnaires.

If written questionnaires are chosen as a method for eliciting data from blind or visually impaired people it is essential that respondents can both read the questionnaire and respond in their preferred reading and writing formats. This may mean providing questionnaires in large print, on coloured paper, in Braille, on audiotape, on disk or in other formats. Care must be taken to ensure consistency across all formats and that the instructions for responding are adapted to suit the format in use. For example, asking respondents to "tick a box" may be fine in print but if the same instruction is given on an audiotape it becomes meaningless. Responses are unlikely to be obtained from people who cannot read the questionnaire or who cannot fill it in. Making questionnaires accessible, therefore, is a crucial part of the research design.

9.3.1.2 Interviews

Interviews involve asking questions usually face-to-face or by telephone. They may be structured, where the interviewer will follow a predetermined series of questions, or unstructured, where the interviewer will ask questions but not necessarily predetermined ones. In many situations interviews will be semi-structured, with a mix of predetermined and spur of the moment questions. Interviews have a number of advantages. In particular, they increase the interaction between researcher and participant and, in doing so, are well suited to the collection of in depth information from individuals. This may be especially important in gathering user requirements for electronic travel aids because a positive interaction between researcher and participant will encourage the participant to discuss things that they may feel uneasy discussing, such as difficulties they have in getting out and about alone.

However, like questionnaires, interviews to elicit user requirements for electronic travel aids may be difficult. They require participants to think about a travel situation whilst sitting still being interviewed. The two situations are very different and participants may have trouble in imagining what they do need whilst travelling.

9.3.1.3 Focus Groups

Focus groups involve discussions between groups of people, with a facilitator. Focus groups not only allow for interaction between researcher and participants but, importantly, they allow for interaction between participants. This opportunity for the participants to interact, to exchange ideas and stimulate each other's thoughts can lead to a much richer source of data than questionnaires or interviews. This rich source of data can be gathered in relatively short time from a number of participants, making them less resource intensive than interviews.

However, focus groups still require participants to think about a mobile situation whilst they are sitting still. Goodman *et al.* (2004) describe the use of visual prompts and scenarios during focus groups to elicit older people's requirements for navigation devices. They demonstrate that visual cues, such as photographs of points along a route, can enable people to think about and discuss a travel situation much better whilst sitting still. However, such visual prompts would be inaccessible to many blind or visually impaired people. Future research needs to consider how non-visual prompts may be used to enhance blind or visually impaired people's ability to think about mobile situations whilst they are stationary.

9.3.1.4 Protocol Analysis

Protocol analysis can be used to identify what information and knowledge is required to perform a task, the nature of mental models used whilst performing a task, the unwritten rules that are applied and the errors that people make and how they rectify them. A number of techniques may be used. An expert (a user) may be recorded thinking aloud whilst they perform a task. A conversation between two or more people may be recorded. A person may be recorded whilst they perform a task and then shown the recording and asked to explain what they were doing and why. A procedural checklist may be used, where predetermined task elements are numbered as they occur. All of these approaches rely on inferences and not on direct data. Misinterpretations are likely to occur, especially if jargon is being used or if either the researcher or participant is making assumptions.

9.3.1.5 Empathic Modelling – A Word of Caution

Empathic modelling, or simulation, is not a user-focused method. This is when a developer or researcher simulates impairment by, for example, wearing glasses that simulate visual impairment. It is easy to think that by simulating impairment a researcher is in a position to understand the needs of a visually impaired person. This is not the case. By simulating a visual impairment the researcher is experiencing what it is like to be suddenly and temporarily without full vision. The researcher is not experiencing what it means to be a visually impaired person, living constantly with and adapting to that impairment and living day in day out with the disabling barriers created by a society that takes little or no account of vision impairment needs. Simulation may help to understand concepts such as tunnel vision but it does not help understand the needs of a person with tunnel vision. Simulation, then, should be used with great caution and should never replace the involvement of visually impaired and blind people themselves.

9.3.2 Developing a User Requirements Specification and Heuristic Evaluation

Eliciting user requirements by itself is not an assessment. These must be drawn up into a requirements specification. Each requirement should be clearly stated in user-centred functional terms. Once a list of user requirements has been established electronic travel aids can be assessed against it to determine how well the aid meets each of the specifications.

Such an assessment can be relatively quick and simple to perform. However, it is based on what users think they need and on the expert evaluator's judgement of how well the device matches those needs. It is not based on their actual experience of using the aid being assessed. Gathering user requirements and performing expert evaluations, therefore, should, at most, form only part of an assessment. They may be particularly useful in the early stages of design, before a testable prototype or product is available. Other methods, giving users hands-on experience with the device are important for a comprehensive assessment.

9.3.3 Hands-on Assessments

Hands-on assessments give users an opportunity to gain experience of using a device. Data collected through hands-on assessments will more accurately reflect what users actually need and want, not just what they think they might need and want.

9.3.3.1 User Trials

User trials give potential users of an electronic travel aid an opportunity to gain hands-on experience of using it within a controlled setting. The performance of the device and of the user can be assessed; the user may be observed and asked questions to elicit information about their experience of using the device. User trials have the advantage of giving users hands-on experience of using the device whilst giving the researcher control over the situations in which the user is using the device.

9.3.3.2 Field Trials

Field trials allow users to gain hands-on experience of using an electronic travel aid in their everyday lives. This is important because of the variety of tasks and situations users of electronic travel aids will encounter and because of the importance of considering the impact of the environment. However collecting data from field trials can be resource intensive and objective data may be hard to acquire.

9.3.4 Methodology Used for Assessments in this Chapter

Assessments of electronic travel aids, in this chapter, are based on the combined results of two research projects. These results were used to generate end-user requirements for electronic travel aids. A range of modern-day aids are assessed against these user requirements. In addition, the results from these research projects have been used to identify training requirements that users have for electronic travel aids. Note that this is not a comprehensive assessment. It involves only minimal hands-on assessment.

9.3.4.1 Project 1: User Requirements for a Navigation System to be Used by Blind and Visually Impaired Pedestrians

This project (Ball 2001) set out to identify the requirements that blind and visually impaired pedestrians have for a navigation system. Though specifically aimed at navigation aids, not mobility aids, the project did, in fact, ascertain many requirements that this group of people have for both navigation and mobility aids. This was, in the main, because study participants did not make a distinction between mobility and navigation.

A short questionnaire was created to explore current navigation strategies of blind and visually impaired pedestrians, the information they felt a navigation system would need to provide and how this information should be presented. This questionnaire was distributed by email. The advantages and disadvantages of using email will be discussed in Section 9.3.4.3. Though the questionnaire asked specifically about navigation, many respondents included information relating to mobility as well. The results of the questionnaires were used for two purposes. They were fed directly into the identification of user requirements. They were also used to determine important issues that needed to be addressed in the next stage of the project, which involved user trials of a navigation system.

A series of eight user trials of the GPS-Talk system (the forerunner of the BrailleNote GPS) was conducted with blind and visually impaired travellers. Each user trial was divided into three parts. First, users were trained to use the system and given ample opportunity to explore the digital maps. Second, each user undertook a trial walk with the system. Due to technical problems with the system some simulation was needed on all occasions (see Figure 9.1). If the system failed, a human provided the information that the system would ordinarily provide. Finally, each user was interviewed about their experience of using the system, how it could be improved and what they would want and need from a navigation system. The results from these user trials informed the identification of user requirements.

9.3.4.2 Project 2: Orientation and Mobility Training

The second project (Ball 2006) aims to identify important issues in orientation and mobility training of blind and visually impaired people and to look at some of the barriers to effective training. The first phase of this has been to interview blind and visually impaired people about their experiences of travelling, the challenges they face, how their travel could be enhanced and about their experiences of learning to travel independently. Rehabilitation workers, providing orientation and mobility training to blind and visually impaired adults have also been interviewed about their experience of providing training. To date, the majority of these interviews have been conducted by e-mail. The results of these interviews have identified a number of issues in orientation and mobility training, some of which apply specifically to training in the use of electronic travel aids and some of which apply equally to all forms of orientation and mobility training. Second, data from these interviews have been combined with data from the first project described

Figure 9.1. GPS-Talk: will it work?

in Section 9.3.4.1, to inform the identification of user requirements for electronic travel aids.

9.3.4.3 The Choice of Methodology

Both of the above projects included blind and visually impaired participants, some of whom had additional impairments. Any user-focused assessment of electronic travel aids will, at some stage, involve blind or visually impaired people. It is therefore worth considering how sensory impairments influenced the research methodologies used in the above studies, as an example of the unique issues that may arise in assessments. Here only a brief summary of the issues is presented. For a more detailed discussion see Ball (2004).

All participants need fair and equal access to information prior to and during the research. Information, therefore, must be provided in a range of reading formats: standard print, large print, coloured paper, audiotape, computer disk, email and Braille were all made available. Electronic information, and indeed the email questionnaires and interviews, were tested for ease of use with screen readers and screen magnification software. Other formats were all proof read and checked for quality against each other. This ensured that high quality information was available in the participants' preferred reading format. If a participant preferred, the information was read to them in full. All participants, therefore, irrespective of their quality of vision, had access to the same information.

Both projects included the use of e-mail to gather data from participants. In the case of the first project the primary reasons for this were that it allowed maximum amounts of data to be collected in a minimum of time with the minimum use of resources. These factors were less of an issue in the second project, where alternatives were offered but the majority of participants chose email. E-mail has a number of advantages and disadvantages for interview-type data collection.

Advantages of e-mail data collection

- Large amounts of data can be collected quickly with a minimum use of resources.
- People can participate at a time and location that is convenient for them.
- It eliminates the need for participants or researchers to travel for face-to-face interviews.
- It allows participants to reflect on their responses before returning them—something that Vernon (1997) identifies as important when conducting disability research on topics which may be sensitive.
- It is a format that is accessible to many potential participants and, in the case of the second project, it is a format that is accessible to the deafblind researcher.

Disadvantages of e-mail data collection

- There is less interaction between researcher and participant than there is in other methods, such as face-to-face or telephone interviews. This results in less opportunity to pursue issues raised by participants and to ask additional questions to elicit further information. To offset this disadvantage, participants can be encouraged to ask questions of the researcher and the researcher can follow up an initial email response with further questions.
- Only the verbal content of communications are transmitted by e-mail. Non-verbal content, such as tone of voice or facial expression, is inevitably lost. Non-verbal communication often conveys important, though in many cases, subtle information about such things as a users emotions and thus is significant in determining what is really meant. However, because e-mail is asynchronous, this is not a major problem, since, for example, nonverbal communication is not needed to indicate turn-taking. When we receive an e-mail we know it is our turn to respond. Whereas, with face-to-face communication, when we receive a piece of information it may or may not be our turn to respond. Whether it is our turn or not is determined by non-verbal communication, such as pauses, questioning looks and eye contact, amongst other things.
- Not all potential participants have access to email or feel comfortable using it. Therefore, to ensure that a representative sample is reached other methods must be employed in addition.

9.3.4.3.1 Validity and E-mail

The validity of data collected by e-mail should be questioned and checked by comparing it to data collected through other means. E-mail has the disadvantage of giving participants time to consider how they think they ought to answer a question and may not elicit genuine responses. For example, in the second project (see Section 9.3.4.2), on two occasions the researcher recognised that participants were quoting from textbooks and not giving their own responses.

A second problem affecting validity also arose in the second project, which, in fact, must be borne in mind in all user-focused assessments of electronic travel aids. With e-mail interviews it is not obvious to the participants whether the actual researcher is disabled or what the nature of the researcher's impairment is. In one case, a participant contacted the researcher to request that some of their responses be changed, after they realised that the researcher was blind. It is not clear why these changes were requested or how widespread this problem is. However, it does raise questions about validity of data. When involving blind or visually impaired people in research it must be borne in mind how the individual researcher may themselves inadvertently influence the responses of participants, even by something as unavoidable as their own impairments.

9.3.4.4 Other Methods for Assessments

Establishing a set of user requirements and using these as the basis for an expert evaluation of electronic travel aids is only the beginnings of a comprehensive user-focused assessment. More robust and comprehensive methods need to be employed if any assessment is to be thorough. Few blind or visually impaired people have direct experience of any kind of electronic travel aid and hands-on experience for users, during assessment, is an important element. User trials, such as those used in the first project described above, where users can test a device within a controlled setting, and field trials, where users can test a device within their everyday lives, are essential for a detailed user-focused assessment.

9.3.4.5 Flaws in the Assessments Presented Here

It should be borne in mind throughout this chapter that the assessments presented here are based on heuristic evaluations against a set of user requirements. These requirements were generated through consultation with blind and visually impaired people. However, this does not constitute a comprehensive user-focused assessment. Keep in mind that what is presented here is only a beginning and is intended to stimulate further user-focused research, where the needs of users, the tasks they need to perform and the context in which they operate are primary considerations to which technology should be adapted.

9.4 Modern-day Electronic Travel Aids

The electronic travel aids discussed in this chapter are a selection of those currently available. They cover both primary and secondary travel aids and both mobility and navigation aids. They utilise a range of technologies—pulse laser and sonar, sweeping sonar and optics. They also make use of a range of different outputs—vibratory and simple and complex auditory. The discussion considers only those electronic travel aids that are carried by the user and require no specific modifications to the environment, though both navigation devices discussed here do use satellite signals.

9.4.1 The Distinction Between Mobility and Navigation Aids

Blind and visually impaired people face two types of challenges when on the move. One is to identify a safe path avoiding and negotiating obstacles and hazards. Such obstacles and hazards may include stationary objects such as posts, walls, signs, overhanging branches, stationary vehicles, moving objects such as pedestrians or moving vehicles, and other hazards such as steps, platform edges or icy patches. This is the challenge of mobility. Aids that assist with detecting, avoiding, negotiating and identifying obstacles and other hazards can be categorised as mobility aids. They assist people in moving safely.

The second challenge faced by blind or visually impaired travellers is that of knowing the current location and establishing how to get from the current location to a destination. This is the challenge of navigation or way-finding. Aids that assist with identifying the current location or a route from one location to another can be categorised as navigation aids—they assist people in finding their way.

9.4.2 The Distinction Between Primary and Secondary Aids

Travel aids can also be categorised as either primary or secondary aids. Primary aids are those which can safely be used alone; they provide sufficient information for the blind or visually impaired traveller to move around independently. The widely used long cane and guide dog are examples of primary aids. So, too, are the UltraCane, LaserCane, and the BAT “K” Sonar Cane, which are discussed later in this chapter.

Secondary aids are those that do not by themselves provide sufficient information for a blind or visually impaired person to safely and independently get around. They must be used in conjunction with a primary aid. They enhance rather than replace the information provided by a primary aid. Secondary aids include the Miniguide, the vOICe, the BrailleNote GPS and the Victor Trekker, which are discussed later in this chapter.

9.4.3 User Requirements: Mobility and Navigation Aids

Most blind travellers, just like most sighted travellers, want to be able to get from one place to another safely, easily, efficiently and with dignity. This wish for a convenient and bother-free journey underlies most of the requirements users have for mobility and navigation aids. It should influence all aspects of the design of mobility and navigation aids, including their appearance, how they output information, how they are operated, and the functions they provide.

In this section we will discuss some of the key user requirements for mobility and navigation aids. This is by no means a comprehensive examination of user requirements, as further research is needed and many requirements will be different for specific devices, tasks, environments and users. Rather, it is presented as a summary of core user requirements and to make explicit the key criteria against which the electronic travel aids discussed later in this chapter have been assessed.

Many requirements apply to all types of electronic travel aids—mobility or navigation, primary or secondary. Few users make a distinction between types of travel aids. For these reasons, this section also makes little distinction. Not surprisingly, requirements relating to appearance, output and operation, for the most part, apply to all types of travel aids and greater distinction between types of devices is seen in the functions that they serve.

Appearance and physical characteristics The prime issue for users relating to the physical characteristics and appearance of travel aids is that they should be unobtrusive, inconspicuous and easy to carry. They want devices that are discrete, do not draw attention from others, are not alienating and are not visibly expensive high tech. For convenience of carrying users want devices that are small and light weight and consisting of a single unit. Since travel aids will be used outdoors in all weathers and where accidents are relatively likely, users also feel that an essential criterion is that they must be extremely robust, able to withstand all weather conditions and able to withstand knocks and falls.

Operation Users need devices that are simple to operate. It must be quick and easy to change modes or issue commands. This is of particular importance for those with additional impairments. For example, those with limited dexterity require buttons that are easy to use and for those with cognitive impairments the operation must be easy to remember and, where appropriate, prompts must be provided. Accidental issuing of commands should be guarded against, for example by a key lock.

One aspect of operation that is of particular importance to many users is the ability to easily and instantly silence the travel aid. This is especially significant as it ensures that the user can speedily attend to other crucial sensory information, such as a public announcement or somebody speaking to them.

It is also important that users can operate the device in a real travel setting where they may have other things to carry or hands may be cold, limiting dexterity. Many users consider hands-free operation extremely useful.

One area of user requirements, relating to operation, which needs further research is that many users wish to be able to choose between different methods of operating a device. For example, a user may prefer the discrete and quiet option of pressing buttons whilst amongst other people but may prefer to use an automatic speech recognition system when away from others and perhaps in environments where their hands may be otherwise occupied or wearing gloves to protect them from cold.

Simplicity of operation applies not only to normal use but also to maintenance. The user must be able to change batteries easily, for example. Similarly, if the system relies on a database, such as is the case with navigation aids, it must be simple for the user to update that database.

Output How the device outputs information to the user is of great importance. The user must be able to make sense of the information provided by the travel aid. User control and choice over output is vital. This should be no surprise. Consider the many diverse situations in which a travel aid may be used—a quiet, obstruction-free country lane, a hushed, cluttered library, a crowded shopping street or alongside a busy road. Users need to be able to adjust the volume and intensity of output to suit the situation and their individual needs. Many users would also like to be able to choose between different modes of output. For example being able to select between auditory or tactile output.

Choice and control are also key in what information is output from the device. Users need to be able to select what information they are given. This is especially true for the more complex devices, navigation aids in particular, where the aid can provide huge amounts of information. Users want to be able to select to be given only certain information. For example if using a navigation aid on a familiar route they may wish to choose only to be told when they reach a point where they have to turn, whereas, on an unfamiliar route they may wish to be given much more information about where they are and what they are passing. Related to this is the need for information to be prioritised. Being warned of an impending collision may be of higher priority than being told what the object is. Being told to turn left at the next intersection may be of higher priority than being told that the user is passing a side street on the opposite side of the road. Similarly, for some users, being told the location of ramps, for example, is essential whereas other users will not require this information. Users want to be able to set their own priorities.

Function Most users consider reliability, accuracy and consistency of information vital. Many users, in fact, consider consistency to be more important than the specific information provided. If the device informs the user of the location of objects or landmarks, it must provide the equivalent information, at the equivalent time, for all objects and landmarks. If this cannot be achieved, some users would prefer to have no information about any objects or landmarks than to have information about some but not about others or to have information at inconsistent times. It is also vital to many users that weather conditions do not interfere with the functioning of the travel aid.

A distinction can be made between two groups of mobility aids. One group provides information only about the presence or absence of an object and its approximate distance (or time until impact). This group includes the Miniguide, Ultracane and Laser Cane. The second group provides richer spatial information, enabling users to perceive the form of objects and spatial relationships. This group includes the BAT "K" Sonar Cane and the vOICe software. Further research is needed to investigate the relative effectiveness of these two groups of devices in different settings and with different users. Dodds (1988) describes how the psychologist Alfred Leonard believed that users of the Sonar Torch (from which the BAT "K" Sonar Cane has developed) were filtering out much of the information it provided. Leonard believed that the complex spatial information was largely being ignored and that users were using it as a yes-no, safe-not safe device. This view is fiercely contested by its inventor, Leslie Kay, whose research shows the complex spatial information is understood and used by users (Kay 1980, 2000). Both these groups of mobility aids enable a user to make judgements about whether they are safe to continue on their planned path. Those providing richer spatial information assist the user in identifying objects. This means that users can better use objects as landmarks. Takamiya and Hamada (1998) emphasise the importance of landmarks in enabling blind and visually impaired people to navigate on both familiar and unfamiliar routes. This suggests that object identification will assist blind and visually impaired travellers. Are all users able to process and act upon this rich source of spatial information? Is this information needed for independent travel, in all situations? Is it of greatest use in certain situations? These questions must yet be answered by objective research in real-world settings.

Navigation aids, combining geographic information systems and global positioning systems, to give blind and visually impaired people access to maps and to real-time positioning information, attempt to address the significant challenge blind and visually impaired people face in navigating. They provide blind and visually impaired people with access to map information that is equivalent to that available to sighted people. This is a very positive step. However, if we take a closer look at the needs of blind and visually impaired people we find that the navigation information they need or would like goes beyond that needed or wanted by sighted people.

The MoBIC project (Petrie and Johnson 1995; Petrie *et al.* 1995, 1996, 1997; Petrie 1996; Gill, 1997) used a combination of interviews, user trials and field trials to establish user requirements for a navigation system and to evaluate a navigation aid. The results of the MoBIC project (Strothotte *et al.* 1995, 1996) and the results of the first project describe in Section 9.3.4.1 show that blind and visually impaired travellers require navigation information specifically designed for their needs. Broadly speaking the additional information required by blind and visually impaired pedestrians falls into two overlapping categories—additional route information and landmarks.

First, blind and visually impaired travellers need extra information that assists them in identifying and following a safe route. This may, for example, include information such as the location of pedestrian crossings, whether there are shared pedestrian and cycle ways and whether or not there are pavements. For some users, it is essential that they are able to identify wheelchair accessible routes. It is of limited help if a navigation aid, for example, tells a blind or visually impaired traveller to turn down a road if in order to do so they must cross a road with railings in the middle of it to prevent people crossing. Although those with good mobility skills may be able to overcome such difficulties, those who are less able and less confident travellers may not and may be discouraged by them.

Second, there is information about landmarks. This may include information about direction and angle of gradients, floor surfacing or fixed obstacles such as telephone boxes. Note that this type of landmark is slightly different to those typically used by sighted people. Sighted people may use building such as pubs or petrol stations as landmarks. These are ineffective as landmarks for blind or visually impaired travellers as they may be unaware of them or have difficulty identifying them. Landmarks for blind or visually impaired travellers must be immediately identifiable and within their awareness. Landmark information serves a number of purposes including a means of validating positioning information, thereby giving the user greater confidence, and enabling a user to continue following a planned route even if satellite signals are unavailable, for example if they have been blocked by tall buildings. Without this additional information blind and visually impaired people may find themselves unable to follow a planned route, arriving close to their destination but being unable to find the entrance or being unable to find their way if positioning information is temporarily unavailable. As pointed out by Golledge *et al.* (1998), a system that provides information about landmarks is enriching the user's experience of the environment. If this information is not provided the device is simply acting as path following aid. An additional important factor is that this extra information can boost a person's confidence; it provides a means for them to check that they are on the right path. This becomes particularly important when we recall the findings of Yerassimou (2002) that fear of becoming lost or disorientated deter many blind or visually impaired people from travelling independently. If navigation aids are to be used by this group of people they must build the user's confidence.

There is still much research to be done to identify user requirements for electronic travel aids. Those requirements presented here are just a beginning. We can, however, see some common themes running throughout all of the requirements discussed here. These include the wish for safe, hassle-free, dignified independent travel. This leads to a desire for travel aids that are discrete, simple and flexible and which allow the user choice and control.

Let us now consider a representative selection of electronic travel aids that are available today. The devices discussed here are chosen as examples of electronic travel aids using different types of technologies and serving different purposes.

The first five devices, the Miniguide, the Ultracane, the Laser Cane, the BAT "K" Sonar Cane and the vOICe, are examples of mobility aids. These are aids to assist the user in avoiding hazards. Two of the above devices, the Miniguide and the vOICe are secondary aids. They augment the information provided by a primary aid, such as cane or guide dog, they do not replace it. The other three mobility aids, the Ultracane, Laser Cane and the BAT "K" Sonar Cane, are examples of electronic primary mobility aids. They provide the essential information for moving safely through the environment; replacing the information provided by other primary aids such as the long cane or guide dog.

The final two devices, the BrailleNote GPS and the Victor Trekker, are examples of navigation aids. These assist users in identifying their location within the macro environment and finding routes between places. These are secondary aids. They provide different information to that provided by a primary aid and must always be used in conjunction with a primary aid such as the long cane or guide dog.

9.4.4 Mobility Aids

Mobility aids are those that assist the user in identifying a safe path through the immediate environment.

9.4.4.1 The Miniguide

The Miniguide is a truly mini, lightweight and versatile secondary mobility aid. It has two transducers. One emits a short pulse of ultrasound. The ultrasound is reflected back off objects in its path. This reflected ultrasound is received back

Figure 9.2. The Audio Miniguide (© GDP Research http://www.gdp-research.com.au. Used with permission.)

Figure 9.3. The Tactile Miniguide (© GDP Research http://www.gdp-research.com.au. Used with permission.)

by the second transducer. The delay between transmitting and receiving back the ultrasound is used to approximate the distance to the object.

Initially there were two versions of the Miniguide. One provided audio output (see Figure 9.2) and the other provided vibratory output (see Figure 9.3). These have now been combined so that one device can be used to provide either audio or tactile output (see Chapter 5). This gives the user greater choice and flexibility. The audio mode beeps to inform the user of objects within the user-selectable range. The distance to an object is represented by the pitch of the beep. High-pitched beeps indicate a close up object. The further away the object is, the lower the beep. The tactile mode uses vibrations to alert the user to objects within range. A distant object results in a slow rate of vibration and a nearby object causes a rapid vibration.

The Miniguide has a number of modes of operation, making it well suited to many different environments. The user can select an object detection range of 8, 4, 2, 1 or 0.5 m. In addition there are modes with reduced sensitivity designed for the special purpose of locating gaps, such as open doors. The audio modes offer a choice of "chirp" or "sweep" output. In "chirp" mode, the device makes a short sound once in each cycle. In "sweep" mode the audio output is continuous. This continuous output assists in the detection of small changes. However, for some users it makes the interpretation of the output too complex. Switching the Miniguide on and off and changing modes is achieved through simple operation of a single button.

The Miniguide performs well against user requirements. In particular, it is small and lightweight, is very versatile, has excellent battery life and is relatively low cost.

Size and weight are important physical characteristics for convenience of use and carrying whilst not in use and making the device inconspicuous—something that many users consider an essential criterion. The Miniguide is usually hand-held but some users have successfully mounted it to a cane, wheelchair or even a hat. This flexibility in where to hold the Miniguide, combined with its multiple and easily user-selectable modes make the Miniguide well suited to many different environments both indoor and outdoor. Hill and Black (2003) has reported an evaluation of the Miniguide based on field trials. One of the most striking findings is the variety of purposes users put their Miniguides to. As well as the typical mobility related obstacle avoidance uses included locating the clothesline, following queues and locating cashiers behind counters. Its flexibility and versatility is one of the Miniguides greatest strengths.

However, it is not without its shortcomings, primarily due to the use of short pulses of ultrasound. As these problems apply equally to the Miniguide and the Ultracane, let us look first at the Ultracane and discuss the disadvantages of both together (see Section 9.4.4.3).

9.4.4.2 The Ultracane

The Ultracane is a primary mobility aid, which combines the long cane with ultrasonic sensors. Two ultrasonic sensors are built into the shaft of a cane. One is positioned to detect obstacles straight ahead, the second to detect obstacles ahead and upwards. Short pulses of ultrasound are transmitted. The ultrasound reflects back off objects within its path and is detected by the two sensors. The delay between transmission and detection provides a basis for estimating the distance to the object. The user is informed of the presence of an object through two vibrating buttons on the handle of the cane—one button for each sensor. This means that a user can distinguish between an object close to ground level and one closer to head level. If a distant object is detected the vibration rate is slow. As the distance to the object decreases, the rate of vibration increases. The user can switch between an object detection range of 3 or 1.5 m, intended for use outdoors and indoors respectively. For further details on the Ultracane see Chapter 6.

The Ultracane has the positive points of providing more information than is provided by a standard long cane whilst being a single unit very similar to the traditional cane. This means that users gain from enhanced mobility information, whilst only needing one hand for mobility purposes leaving the other free to, for example, carry bags. Having at least one hand free is something many users consider important. It offers some flexibility, with user selectable ranges, again something of considerable importance. It also has the advantage of making use of the long cane, something that most users will be familiar with before using the Ultracane, and so is fairly straightforward to learn to use.

Though the Ultracane is based on the traditional long cane, it is a little heavier and the handle a little thicker than most long canes. This results in some users finding it difficult to maintain their usual cane technique. This could have implications for the user's ability to detect drop-offs, such as the edges of platforms, using the cane. This is because both the two-point touch technique and the constant contact techniques, when properly used, restrict the directions in which the wrist moves to just side to side movements and when the cane tip drops over the edge of a drop-off, even a very small one, this results in a subtle up-down stretch in the user's wrist, alerting them to the change of level (Wall 2002). If correct cane technique is not maintained, the users wrist may be allowed to move up and down, as well as side to side, making the stretch as the cane tip descends over an edge less noticeable. Other problems relate to the use of short pulses of ultrasound.

9.4.4.3 Ultrasonic Mobility Aids: Some Disadvantages

Both the Miniguide and Ultracane use short pulses of ultrasound to detect objects within range. There are some disadvantages with this. Note that here we are not talking just of the weaknesses of the technology but how they impact upon users. Assessments of electronic travel aids must always consider how the user may be affected by the technology.

Different surfaces differ in how well they reflect ultrasound. Generally, hard surfaces such as stone or metal reflect ultrasound well, whereas, soft objects such as bushes or soft furnishings absorb some of the ultrasound so reflect less well. This means that a hard object, one that reflects lots of ultrasound, will be detected at a greater distance than a soft one, which reflects less ultrasound. There is, then, an inconsistency in the amount of warning users are given about the presence of an object. Consistency of information provided by an electronic travel aid is an important user requirement.

Ultrasonic mobility aids are subject to interference from environmental sources of ultrasound. For example, some machinery and the air brakes on buses and lorries can emit ultrasound. If this is picked up by an ultrasonic mobility aid, such as the Miniguide or Ultracane, the aid will react as if an object has been detected. The user will wrongly be alerted to an object, when in fact none is present. This may result in confusion. In practice, however, interference is not common and, as users gain experience of using a device, they will come to suspect when interference may be giving rise to false alarms.

Ultrasonic mobility aids can detect objects above ground level. However, they cannot satisfactorily detect drop-offs, such as steps going down or the edges of platforms. This type of hazard can be particularly significant and so the ability to reliably detect drop-offs is essential. Note that the Miniguide should always be used in conjunction with a primary aid and that it does not interfere with the use of a primary aid. The user should therefore choose a primary aid that can detect drop-offs. The Ultracane, on the other hand, is a primary aid and, though if correct cane technique is used the user will be able to detect drop-offs through the cane itself (note the cane not the ultrasonic detectors) but if cane technique is not correct drop-offs may be missed.

Finally, mobility aids using short pulses of ultrasound, such as the Miniguide and Ultracane, provide the user only with information about the nearest object. The aid will always react to the closest object. If there is more than one object within range the user will only be warned of the nearest one and will remain unaware of the others. For many situations this is adequate. It tells the user whether or not the way is clear for them to take the next steps, so achieves the most essential element of safe movement. An important user requirement is that information should be prioritised. This is a form of prioritisation. The user, if they continue on the same path, will collide with the nearest object first, so this can be considered as priority. However, in many situations more information, such as the size and shape of the object and the spatial relationship between one object and another are important for enabling the user to make decisions about how to negotiate an obstacle or to allow the user to identify the object in order that they may use it as a landmark.

Other mobility aids exist which overcome at least some of these problems of short pulse-based ultrasonic devices.

9.4.4.4 The Laser Cane

The Laser Cane, available from Nurion Industries, is a primary mobility aid that combines a long cane with a laser obstacle detection system. Short pulses of laser are transmitted. The laser light reflects off obstacles in its path. This reflected light is detected by the sensors mounted on the cane. Sensors are positioned to detect obstacles straight ahead near ground level and at head height, as well as obstacles to either side and a downward pointing sensor is used to detect drop-offs. Information about the presence of obstacles is provided to the user through both vibrations and beeps. The user may select to switch off the beeps for silent operation.

The Laser Cane has two major advantages over the Miniguide and Ultracane, relating to the use of laser rather than ultrasound. First, laser is able to detect drop-offs, giving the user extra warning of these hazards, something which is important for safe travel. Second, laser is, largely, immune to interference. It is extremely unlikely indeed that users will be falsely warned of an object (see Dodds 1988, for a discussion of these issues).

The Laser Cane, however, does have its own disadvantages. One important disadvantage is that not all materials will be detected. For example, light will pass straight through transparent glass, it will not be reflected, so the laser sensors will not detect it. The user will continue to walk until the cane itself hits against the glass object. So, like sonar aids, there is an inconsistency in the warning users will be given about the presence of obstacles.

A further disadvantage with the Laser Cane, due to its use of short pulses of laser, is that, like devices using short pulses of sonar, it only provides the user with information about the nearest object. The user is informed only of the approximate distance to the closest object. Multiple objects are ignored and information about spatial relationships between objects is unavailable to the user. Again, for many situations this is adequate and can be thought of as a form of prioritisation of information. However there are some situations where more spatial information is required. Let us now look at two devices that provide a much richer source of spatial information.

9.4.4.5 The BAT "K" Sonar Cane

The BAT "K" Sonar Cane, from Bay Advanced Technologies, is a primary mobility aid that combines the long cane with an advanced sweeping sonar system (see Figure 9.4). It uses a modern revision of one of the earliest electronic travel aids, the Sonar Torch, to provide advanced spatial information. It has two transducers. One emits a prolonged beam of ultrasound. The ultrasound is reflected off obstacles in its path and this reflected ultrasound is detected by the second transducer. Distance to an object is represented to the user through beeps of differing pitches. The further away the object is, the higher the beep. As the cane is moved from side to side this sonar beam sweeps across with it. This results in a continuous and unique pattern of beeps being created as the beam sweeps across each object or across multiple objects. This unique pattern of beeps is known as a "sound signature". By remembering these sound signatures users can learn to identify objects and gather detailed spatial information about the surroundings. This enables travellers to identify objects as landmarks, as well as to identify an optimal route around or between obstacles. A further advantage of the BAT "K" Sonar Cane is that the sonar unit can be easily detached from the cane and used independently of it, providing for increased flexibility and meaning that damage to the cane means only the cane need be replaced (see Figure 9.5). The sonar torch may be used alone but the cane provides additional safety in cases where the sonar may miss objects or drop-offs. The Sonar Torch has a maximum range of 5 m and the user may choose between long or short range.

Like other sonar devices, the BAT "K" Sonar Cane is occasionally subject to interference from environmental sources of ultrasound, such as the air brakes on buses, and some surfaces will be detected at greater distance than others.

Users must interpret the complex auditory information from the BAT "K" Sonar Cane. Further research is needed to establish what impact this perceptual demand

Figure 9.4. The BAT "K" Sonar Cane (© Bay Advanced Technologies http://www.batforblind.co.nz. Used with permission.)

Figure 9.5. The Sonar Torch detached from the BAT "K" Sonar Cane (© Bay Advanced Technologies http://www.batforblind.co.nz. Used with permission.)

Figure 9.6. A typical set-up for the vOICe (© Barbara Schweizer and used with permission. Photo kindly donated by Peter Meijer of The vOICe website http://www.seeingwithsound.com.)

may have upon the user's ability to attend to and process other sensory information and how this may impact upon their travelling ability.

9.4.4.6 The vOICe

The vOICe is software which claims to substitute sound for low-vision. The software can be run on a PC or notebook with a camera or even on some modern camera mobile phones. The camera takes images of the environment; these are converted into sound pictures. Visual contours are represented by pitch of beep. This results in every object and scene giving a unique pattern of beeps, its own sound signature. By remembering these sound signatures and learning to interpret the patterns of sounds users can learn to visualise what is around them (see Figure 9.6). For experienced users of the vOICe, it is equivalent to having a small amount of vision. It can, therefore, be used as a secondary mobility aid. It has the key advantages of distance viewing and of being free software that can be run on readily available hardware.

The vOICe software can be run on readily available hardware. How conspicuous or inconspicuous the system is will, therefore, depend on the hardware being used.

The vOICe, like the BAT "K" Sonar Cane, provides complex auditory information that must be processed by the user. This places additional cognitive demands upon the user. Further research is needed to establish the consequences of this on the user's attention and travel ability.

9.4.5 Mobility Aids: Have They Solved the Mobility Challenge?

There is little or no doubt that there are a number of mobility aids available today that perform well in detecting, and in some cases identifying, objects. So do these devices solve the challenge of mobility, of finding a safe path through the immediate environment, for blind and visually impaired people?

9.4.5.1 Appearance

Most modern mobility aids are fairly inconspicuous, though there are some exceptions. This is to say that most are no more noticeable or aesthetically alienating than a long cane. This said, it should be pointed out that some visually impaired people find the conspicuousness of a long cane or guide dog unacceptable. Some individuals will not use any visible mobility aid, preferring to attempt to pass as sighted without the perceived labelling that goes with using a cane or guide dog. Being discreet and unobtrusive is important for many users, even those who accept and welcome the use of a mobility aid.

9.4.5.2 Operation

The handling and operation of most modern mobility aids is simple. Most allow users to switch the device on and off quickly and easily. Being able to silence a mobility aid instantly is of great importance as it allows the user to attend immediately to other crucial sensory information, such as a public announcement or somebody speaking to them. Most modern mobility aids also allow users to change quickly between different modes. In the rapidly changing travel environment this is vital. Imagine, for example, a user coming into a crowded building from a quiet and uncluttered street. Outdoors on the empty street they may walk at a brisk pace, therefore requiring maximum warning of upcoming hazards. Once inside the crowded building a much shorter range will be required. They must be able to make this adjustment without undue pausing. It should, however, be borne in mind that many modern mobility aids require small switches or buttons to be operated. For some potential users, those with limited dexterity, this will be a significant difficulty and may become more significant as miniaturisation of devices continues. Simple operation, by all potential users, is important.

9.4.5.3 Output

Modern mobility aids mostly use vibrations, simple beeps or complex patterns of beeps to inform the user of objects in the environment. All of these methods appear to be effective. It should be noted that most offer only one method of output. Of the five mobility aids discussed in this chapter, only the Laser Cane and the Miniguide allows users to have combined vibrotactile and auditory feedback or to choose between them. Flexibility in output is important to allow users to operate effectively in different environments.

There are currently no tactile mobility aids that assist users in identifying objects. Future developments may see electro-tactile displays used to provide users with tactile information about size, shape and spatial relationships. For a discussion of the use of electro-tactile stimulation as a form of sensory substitution for blind people see Chapter 4 of this volume and the work of Bach-y-Rita *et al.* (1998).

9.4.5.4 Function

Mobility aids can be categorised as primary or secondary aids. Primary aids provide the essential information to enable the user to make the decision as to whether or not it is safe to take the next steps. Secondary aids augment the information provided by primary aids but do not by themselves provide all of the necessary information. Modern mobility aids include both primary and secondary aids. Primary aids have the advantage of (usually) being a single unit and requiring only one hand to operate it. Secondary aids are in themselves (usually) a single unit but must be used in conjunction with a primary aid. The user, therefore, may have a mobility aid in each hand, leaving neither hand free for other things. Most primary mobility aids combine the long cane with an electronic system. What, then, happens if the cane is damaged by, for example, impact with an object or by a moving vehicle? In many cases, if the damage is irreversible, the entire device must be replaced. Secondary aids are less vulnerable to this kind of damage than are primary aids.

Another distinction can be made between two groups of mobility aids—those which simply inform the user of the presence or absence of obstacles, and thus tell the user whether it is safe or unsafe to continue on the present path, and those which provide richer spatial information about the nature of objects and spatial relationships between objects. Further research is needed to establish how much spatial information is required in a variety of situations and whether all users can benefit from rich spatial information.

9.4.6 Navigation Aids

These are devices that assist the user with way-finding.

9.4.6.1 The BrailleNote GPS and the Victor Trekker

The BrailleNote GPS (see Figure 9.7) and the Victor Trekker are both navigation aids. They can provide assistance before setting out to a destination by calculating a route on a digital map and enabling users to explore that route. Once the user has set out, these devices use the global positioning (satellite) system to identify the user's location. This location information can be combined with the information from the digital maps to provide directions or information about the route, such as what points of interest are nearby (see Figure 9.8). The BrailleNote GPS software is an optional extra for the BrailleNote or VoiceNote—notetakers that provide access to a number of applications including wordprocessing, email and organiser. The Victor Trekker is currently available as a dedicated unit, whose only purpose is as a navigation aid (see Figure 9.9).

Figure 9.7. The BrailleNote GPS (© PulseData Europe http://www.pulsedata.com. Used with permission.)

Figure 9.8. The BrailleNote GPS in action (© PulseData Europe http://www.pulsedata.com. Used with permission.)

Figure 9.9. The Victor Trekker

9.4.7 Navigation Aids: Have They Solved the Navigation Challenge?

Navigation aids provide blind and visually impaired people with access to maps and to real-time positioning information that would otherwise be inaccessible to them. So have these devices solved the challenge of navigation, of locating oneself within the macro-environment and planning and following a route through it, for blind and visually impaired people?

9.4.7.1 Appearance

The compact units used for accessible navigation aids, such as the BrailleNote GPS and the Victor Trekker, are fairly discreet although they are visibly high-tech devices. Their appearance is not alienating and with the increasing number of high-tech electronic devices being used by many people, disabled and non-disabled alike, they do not make the user too conspicuous. However, some potential users do have concerns about carrying anything that looks like expensive technology and would prefer that the technology could be completely hidden in a pocket or mundane and ordinary looking bag.

9.4.7.2 Operation

The functions available on navigation aids require the user to input much more to the device than they must with mobility aids. Inevitably this means that their operation is more complex. However, given this increased necessity for user input, the operation of these devices is relatively simple. Currently they require users to be able to operate a keyboard or to press tactile-marked buttons on a touch sensitive screen. Future developments may see the introduction of speech recognition onto these devices. This is a popular option for users as it allows for hands-free operation. However, some users report that they would feel self-conscious speaking into a device and would prefer to have the option to use a keyboard. Simple operation and a user-selectable choice of the mode of input are important issues to consider.

9.4.7.3 Output

Output from navigation aids is through synthesised speech or refreshable Braille. Both can be effective. It should be remembered, however, that refreshable Braille displays are especially vulnerable in wet weather and that synthesised speech may be difficult to hear in noisy settings, such as amidst heavy traffic. In the project described in Section 9.3.4.1, both questionnaire respondents and participants of user trials suggested other possible forms of output including the use of beeps and vibrations, possibly in addition to speech or Braille. Future research is needed to establish the most effective and efficient forms of output and this research should consider the needs of a wide range of users.

9.4.7.4 Function

The positioning system relies upon the GPS receiver being able to pick up GPS satellite signals. If these signals are blocked, as they can be by tall buildings or as they are indoors, the system is unable to update the user's location. This can lead to confusion. So, too, can inaccuracies in the positioning information. Imagine, for example, that a user is approaching a road down which they should turn but there is a driveway shortly before it. If the positioning information is accurate the user will cross over the driveway and turn down the road correctly. If, however, the positioning information is inaccurate, the user may be told to turn too early and will turn into the driveway, only to experience confusion and become lost and disorientated.

Blind and visually impaired travellers require navigation information that goes beyond that needed by sighted travellers. This additional information falls into two broad and overlapping categories—extra route information, such as the location of pedestrian crossings, and information about identifiable landmarks. Such information assists them in following routes, even if satellite signals are temporarily lost and helps to boost their confidence.

Neither the BrailleNote GPS nor the Victor Trekker provides information about landmarks or the very detailed and specific information about routes that many visually impaired and blind travellers need. Users may add information themselves, in the form of a point of interest or by editing the information provided at waypoints. This is a useful facility but does not help if the user is travelling in an area for the first time. Providing such detailed information would require frequent detailed audits. The information would need to be updated very often. It takes little imagination to picture how difficult providing this up-to-the-minute, detailed information, for extensive areas would be.

9.5 Training

Blind or visually impaired individuals need to learn to be safe, confident independent travellers. They must learn concepts of orientation. They must learn to use environmental cues. They must learn to make effective use of residual senses. They must learn to make effective use of their chosen mobility and/or navigation aids. They must also acquire confidence in their ability to travel as a blind or visually impaired person. Effective orientation and mobility training must address all of these issues and more.

Orientation and mobility training for blind and visually impaired people began in the middle of the twentieth century at rehabilitation centres for war-blinded veterans in the US. Orientation instructors taught the veterans concepts and skills such as spatial awareness and echolocation. Dissatisfaction with these techniques led to the development of the long cane and the two-point touch technique. For a discussion of the beginnings of formal orientation and mobility training see Bledsoe (1980). These techniques are still the basis for orientation and mobility today, although they have been refined and adapted. The long cane was introduced into the UK in the 1960s but at first it was not well received (Dodds 1996). Today, most orientation and mobility training is deeply rooted in these practical and skills-based beginnings. Most orientation and mobility training focuses on the teaching of spatial and orientation concepts, the use of environmental cues, the use of residual senses and refinements of the early techniques for safe travel such as protective techniques and the use of the long cane.

What training, then, is available for electronic travel aids? Most, if not all, electronic travel aids come with user instruction manuals. These tell the user how to operate the device. Not all are easily understood by novices who may be confused by technical jargon. More importantly, however, though they deal with the functions and operation of the device, few provide training exercises. One exception is the BAT "K" Sonar Cane, which comes with a series of training exercises to guide the user through learning the concepts involved with using it. Suppliers of the two navigation aids discussed in this chapter, the BrailleNote GPS and the Victor Trekker, can provide training in the use of these devices but again this training focuses on operating the aid and largely ignores other aspects of learning to travel independently.

Yet even when training exercises are provided, the user must transfer what they have learned into the real world if they are to use the device with maximum efficiency. Even if a user succeeds in learning to use a mobility or navigation aid from written material or from a training session this is of limited benefit if they lack the other skills and attributes needed for independent travel. User instructions or a training manual will suffice for some users but many would benefit from training from a suitably qualified trainer, who can address the full range of skills needed for independent travel.

Most orientation and mobility instructors (rehabilitation workers) have only limited knowledge of and experience with electronic travel aids. Furthermore, a substantial number of orientation and mobility instructors are not in favour of electronic travel aids. For some this is because they believe that the more established tools of long cane and guide dog are all that is needed for blind and visually impaired people to travel independently. For others, it is a political issue, the basis for which is beyond the scope of this chapter. The consequence, however, is that few blind or visually impaired people can obtained training in the use of electronic travel aids from a trained and qualified orientation and mobility instructor. Most users of electronic travel aids find that they must train themselves.

Most, possibly all, suppliers and developers of electronic travel aids recommend that blind and visually impaired people should be competent travellers before they start to use their electronic travel aid. Given the lack of training that integrates the use of electronic travel aids into a broader orientation and mobility programme, this caution appears to make sense. However, consider the following questions? If a blind or visually impaired person is already a competent and confident traveller will they be prepared to purchase a (costly) electronic travel aid? Would the benefits of an electronic travel aid, as foreseen by a competent and confident traveller, justify the cost? It seems probable that only those interested in technology may choose to use an electronic travel aid but that others will continue to travel without. At the other end of the scale, people with poor travel skills, those who lack competence and confidence to travel, are advised against the use of an electronic travel aid and so will never benefit from them.

If electronic travel aids are to reach the widest possible range of users, and if they are to be used with greatest efficiency and benefit, it seems that what is needed is a comprehensive approach to training blind and visually impaired people to be independent travellers. This must combine training in the established concepts and skills of orientation and mobility with training in the use of user-chosen electronic travel aids. Importantly, it must also address psychological issues such as motivation and confidence (Dodds 1989; Dodds *et al.* 1993). Training in the use of electronic travel aids should be an integral part of orientation and mobility training programmes and cease to be left to individuals to do the best they can.

9.6 Chapter Summary and Conclusions

Electronic travel aids must be assessed from the user's perspective. This may be achieved in many different ways. The choice of method(s) must be suited to the purpose of the assessment. The user, the tasks he or she needs to perform and the context in which these take place must drive the development of technology.

Modern-day electronic mobility aids use a range of sensing systems including sonar, laser and optics. Some provide basic information about the absence or presence of objects. Others provide complex spatial information such as the shape of individual objects and spatial relationships between objects. They use vibrations or simple or complex auditory output. Each approach has advantages and disadvantages. Further research is needed to identify optimal solutions for a wide range of users in a wide range of settings.

Modern-day navigation aids provide blind and visually impaired people with access to maps and location information that would otherwise be inaccessible to them. They combine geographic information systems with positioning information from the global positioning system. Though making available otherwise inaccessible information blind and visually impaired travellers need and want access to more information than is currently provided. Ways of providing this additional information need to be further developed. Output from navigation aids is through synthesised speech or Braille. Other formats for presenting this information need to be explored.

Currently there is little training available in the use of electronic travel aids. What training there is tends to relate to the operation of the device. It is proposed that an integrated approach to training in the use of electronic travel aids and other aspects of independent travel needs to be developed.

Currently available electronic travel aids all have their advantages and disadvantages. The successes so far indicate that there are substantial potential benefits to be gained through the use of electronic travel aids by blind and visually impaired people. By engaging with users, by undertaking user-focused assessments to identify how devices can be improved and by taking a user-centred approach to design researchers and developers can ensure that the electronic travel aids of tomorrow will enhance the mobility of all blind and visually impaired people. The benefits will be seen most quickly if potential users, developers and orientation and mobility instructors work together to ensure that the needs of users are at the centre of future developments and that training in electronic travel aids becomes an integral part of orientation and mobility programmes. There is still significant work to be done before the challenges of mobility and of navigation for blind and visually impaired people are solved. By working together and taking a user-centred approach, refinements of existing devices and new innovations will overcome these challenges.

Acknowledgement. The author would like to thank GDP Research for permission to include Figures 9.2 and 9.3, Bay Advanced Technology for permission to include Figures 9.4 and 9.5, Peter Meijer for permission to include Figure 9.6 and PulseData Europe for Permission to include Figures 9.7 and 9.8.

Questions

Q.1 Describe the challenges that blind and visually impaired people face when travelling.

Q.2 At what stage(s) in the design process are user-focused formative assessments useful? Why?

Q.3 What is meant by "user-focused assessments"? List at least five methods that can form part of a user-focused assessment. Why is simulation NOT a user-focused method?

Q.4 The Miniguide, Ultracane and the Laser Cane are examples of obstacle detectors (safe/not safe devices). The Bat K Sonar Cane and the vOICe provide rich spatial information. What difference does this make to the user?

Q.5 Successful navigation aids need to provide blind/visually impaired users with more detail than is provided on standard maps. Give at least two examples of the extra information needed and, for each, say why it is needed.

Q.6 What are the advantages of integrating training in the use of electronic travel aids into a broader travel training programme?

Projects

P.1 Choose three of the following methods for identifying user requirements or assessing electronic travel aids. Find out as much as you can about the methods you have chosen and write a list of the advantages and disadvantages of using it and any special considerations that there may be using it with blind or visually impaired people. Choose from: questionnaires; interviews; focus groups; user trials; field trials; heuristic evaluation.

P.2 Devise a questionnaire or structured interview to find out about blind and visually impaired people's experiences of using any kind of travel aid and why they chose the aid(s) they use over other devices. Conduct the questionnaire with as many blind or visually impaired people as you can. Analyse your results to identify the main criteria your respondents used for selecting their travel aids. What does this tell you about the design of electronic travel aids? Take a critical look at your questionnaire. How could it be improved?

P.3 Design a prototype travel aid for use by blind or visually impaired people. Conduct a focus group with between four and eight blind or visually impaired people. Find out about the challenges they face with mobility or navigation. Explore with them how your device fits in with this. From the discussions, how could your device be improved? What mistakes did you make? What did you get right?

P.4 Choose one electronic travel aid, if possible one that you have access to. Find out as much as you can about it. What do you perceive as its strengths and weaknesses? Conduct a focus group with between four and eight blind or visually impaired people. What do they perceive as the strengths and weaknesses of your chosen device? Do they agree with you? If not, can you work out why?

P.5 Choose one primary and one secondary simple obstacle detecting mobility aid (the Miniguide, Ultracane or Laser Cane are discussed in this chapter but there are others you may choose). Find out as much as you can about them. From a user perspective, what are the advantages and disadvantages of each? Compare and contrast them.

P.6 Choose one simple obstacle detecting mobility aid (the Miniguide, Ultracane and Laser Cane are discussed in this chapter but there are others you may choose) and one mobility aid that provides complex spatial information (the BAT "K" Sonar Cane and the vOICe are discussed in this chapter but you may find others you can choose). Find out as much as you can about each. From a user perspective, what are their advantages and disadvantages? Compare and contrast them.

P.7 Find out as much as you can about two navigation devices. The BrailleNote GPS and the Victor Trekker are discussed in this chapter but there are others you may choose. From a user perspective, what are the advantages and disadvantages of each? Compare and contrast them.

P.8 Find out about protocol analysis. Use it to identify the information a blind or visually impaired person uses to navigate a short familiar route. Be aware that talking aloud or your presence may distract the blind or visually impaired person. If this happens offer help but don't insist on helping. Write down all the information they use. By each item, mark whether or not it is provided by either the BrailleNote GPS or the Victor Trekker.

P.9 Working as a pair or as small group, choose one electronic travel aid. Imagine that you have to train a competent blind or visually impaired traveller to use it. What would you include in the training and why? How would this differ if the person you had to train were a poor traveller?

References

Bach-y-Rita, P., Kaczmarek, K.A., Tyler, M.E., and Garcia-Lara, M., 1998, Form perception with a 49-point electrotactile stimulus array on the tongue: A technical note, Journal of Rehabilitation Research and Development, Vol. 35, pp. 427–430

Ball, E.M., 2001, An Investigation of the Information Requirements for a Navigation System to be used by Pedestrians who are Visually Impaired, MSc thesis, Department of Human Sciences, Loughborough University

Ball, E., 2004, Breaking the information and communication barriers. In McCabe, P.T. (ed) Contemporary Ergonomics 2004. CRC Press: Boca Raton, pp. 525–529

Ball, E.M., 2006, Orientation and Mobility Training for Blind and Visually Impaired Adults, PhD thesis (In preparation), the Ergonomics and Safety Research Institute, Loughborough University.

Bledsoe, C.W., 1980, Originators of orientation and mobility training, In R. L. Welsh and B. B. Blasch, (Eds) Foundations of Orientation and Mobility, pp. 581–624, American Foundation for the Blind

Campion, J., Greenhalgh, C., and Knight, J., 2003, Mind the gap, Leonard Cheshire

Clark-Carter, D.D., Howarth, C.I., Heyes, A.D., Dodds, A.D., and Armstrong, J.D., 1981, The Visually Handicapped in the City of Nottingham 1981: a survey of their disabilities, mobility, employment and daily living skills, University of Nottingham

Dodds, A.G., 1988, Mobility Training for Visually Handicapped Adults: a person-centred approach, Croom Helm, London

Dodds, A.G., 1989, Motivation reconsidered: the role of self-efficacy in rehabilitation, British Journal of Visual Impairment, Vol. 1, pp. 11–15

Dodds, A.G., 1996, The way we were: the genesis of training for specialist staff, British Journal of Visual Impairment, Vol. 14, No. 3

Dodds, A.G., Flannigan, H., and Ng, L., 1993, The Nottingham Adjustment Scale: A validation study, International Journal of Rehabilitation Research, Vol. 16, No. 2, pp. 177–184

Gill, J., 1997, Mobility of Blind and Elderly Interacting with Computers, Royal National Institute for the Blind

Golledge, R.G., 1993, Geography and the disabled: a survey with special reference to vision impaired and blind populations, Transactions, Institute of British Geographers, Vol. 18, pp. 63–85

Golledge, R.G., Klatzky, R.L., Loomis, J.M., Speigle, J., and Tietz, J., 1998, A geographical information system for a GPS-based personal guidance system. International Journal of Geographical Information Science, Vol. 12, No. 8, pp. 727–749

Goodman, J., Dickinson, A., and Syme, A. 2004, Gathering requirements for mobile devices using focus groups with older people, Cambridge Workshop on Universal Access and Assistive Technology, Cambridge, UK

Hill, J., and Black, J., 2003, The Miniguide: a new electronic travel device, Journal of Visual Impairment and Blindness, Vol. 97, No. 10, pp. 655–656

Kay, L., 1980, Air sonars with acoustical display of spatial information. In Busnel, R.G. and Fish, J.F. (eds) Animal sonar systems. New York: Plenum Press

Kay, L., 2000, Auditory perception of objects by blind persons using bioacoustic high resolution air sonar, JASA, Vol. 107, pp. 3266–3275

Oppenheim, A.N., 2000, Questionnaire Design, Interviewing and Attitude Measurement, Continuum, London

Petrie, H., 1996, MoBIC: A mobility aid to increase the independence of blind travellers. In Tellevik and Haugum (Eds), Procs. of International Mobility Conference 8, Trondheim, Norway: Tambartun National Resource Centre, Norway

Petrie, H., and Johnson, V., 1995, MoBIC: an aid to increase the mobility of blind and elderly travellers, Proceedings of Third European Conference for the Advancement of Rehabilitation Technology (ECART3), National Secretariat for Rehabilitation, Lisbon

Petrie, H., Johnson, V., Strothotte, T., and Reichert, L., 1995, MoBIC: an aid to increase the mobility of blind and elderly travellers, 7th International Conference on Mobility and Transport for Elderly and Disabled People, University of Reading, U.K.

Petrie, H., Johnson, V., Strothotte, T., Raab, A., Fritz, S., and Michel, R., 1996, MoBIC: designing a travel aid for blind and elderly people, Journal of Navigation, Vol. 49, No. 1, pp. 45–52

Petrie, H., Johnson, V., Strothotte, T., Raab, A., Michel, R., Reichert, L., and Schalt, A., 1997, MoBIC: An aid to increase the independent mobility of blind travellers, The British Journal of Visual Impairment, Vol. 15, No. 2, pp. 63–66

Strothotte, T., Petrie, H., Johnson V., and Reichert, L., 1995, MoBIC: user needs and preliminary design for a mobility aid for blind and elderly travellers. In I. Placencia-Porrero and R. Puig de la Bellacasa, (Eds.), The European context for assistive technology, IOS Press, Amsterdam

Strothotte, T., Fritz, S., Michel, R., Raab, A., Petrie, H., Johnson, V., Reichert, L., and Schalt, A., 1996, Development of dialogue systems for a mobility aid for blind people: initial design and usability testing, Procs. of ASSETS '96: The Second Annual Conference on Assistive Technology, ACM Press, New York, ISBN 0-89791-776-6

Takamiya, S., and Hamada, S., 1998, Information used by visually impaired people while walking, Bicycle and Pedestrian Research, Vol. 1636, pp. 104–109

Vernon, A., 1997, Reflexivity: the dilemmas of researching from the inside, In C. Barnes and G. Mercer, (Eds), Doing Disability Research, Disability Press, Leeds

Wall, R., 2002, Biomechanical substrates of the two-point touch cane technique: a review of research, Journal of Visual Impairment and Blindness, Vol. 96, No. 2

Wilson, J.R., and Corlett, E.N., 1995, Evaluation of Human Work: a practical ergonomics methodology, Taylor and Francis, London

Yerasimuo, N., 2002, Travellers' tales – making journeys safer, Royal National Institute for the Blind, London

Resources

Bay Advanced Technology. The BAT "K" Sonar Cane. http://www.batforblind.co.nz

GDP Research, The Miniguide, http://www.gdp-research.com.au

Meijer, P., The vOICe – Seeing with Sound. http://www.seeingwithsound.com

Nurion Industries. The Laser Cane. http://www.nurion.net

PulseData International. The BrailleNote GPS. http://www.pulsedata.com

Sendero Group. The BrailleNote GPS. http.www.senderogroup.com

Sound Foresight. The Ultracane. http://www.soundforesight.co.uk; http://www.ultracane.com

Visuaide, The Victor Trekker, http://www.visuaide.com

10 Accessible Environments

Learning Objectives

There is an increasing awareness of social responsibility in the community and the benefits to society of social inclusion, including for disabled people. In many countries this awareness is reinforced by anti-discrimination legislation and advisory and regulatory guidelines. However, in many other cases, even where there is legislation, it is not supported by a serious commitment to change. Despite this, there have been some real advances, including in the design of the urban community environment for accessibility. In this chapter, aspects of the accessible environment relevant to visually impaired and blind people are described. The learning objectives for the chapter are as follows:

- Understanding an overall view of the design of accessible environments for visually impaired and blind people.
- Learning about assistive technology systems that are used to make urban environments accessible to visually impaired and blind people.
- Understanding the basic principles of a number of embedded navigational and information system that could be used to improve environmental accessibility.

10.1 Introduction

This chapter is concerned with two main concepts. The first is that of inclusiveness for all members of the community and the second is how to realise this inclusiveness in all aspects of the physical urban environment. Existing legal and administrative mechanisms are discussed in this opening section of this chapter. This is followed by a discussion of specific assistive technology systems and approaches that can be used to create an accessible environment for visually impaired and blind people.

10.1.1 Legislative and Regulatory Framework

Although some progress has been made towards the recognition of the importance of full social inclusion for all members of the community, considerable further work will still be required. In many different countries, there are considerable differences in the extent and type of legislation and how it is implemented, enforced and monitored. For legislation to be successful, the procedures for implementation, enforcement and monitoring are particularly important. However, progress has been rather limited and the impact of potentially useful legislation has often been reduced by the lack of sufficiently stringent sanctions for non-compliance and inadequate attention to monitoring compliance and the impacts of the legislation.

However, despite the drawbacks in the remit of much of the legislation, it has led to the development of regulations and guidelines for accessibility. In some cases, the legislation has included the creation of bodies to oversee the implementation of the law or particular aspects of it. For instance, the UK Disability Discrimination Act 1995 set up a Disability Rights Commission. A number of professional bodies have produced guidelines on accessibility for their members or the profession in general. Such guidelines may be based on ensuring compliance with the legislation or could go beyond it. The physical environment involves several different types of installations and services. For historical and other reasons they may be overseen by a number of different professional bodies, which can complicate implementing and overseeing accessibility. As an illustrative example, a brief discussion of building accessibility legislation in the USA is presented. Although by no means without flaws, environmental accessibility legislation seems more advanced in the USA than in Europe.

10.1.1.1 Building Accessibility Legislation in the U.S.A.

Primary legislation. The first piece of legislation was the Architectural Barriers Act (ABA) of 1968. This covered the accessibility standards required for premises and facilities constructed with Federal funds. This was followed by the most significant piece of US legislation on this subject, namely, the Americans with Disabilities Act (ADA), 1990. This is a civil rights law prohibiting discrimination because of disability. The areas covered in this act are very wide ranging and include employment, state and local government, public accommodation, commercial facilities, transportation and telecommunications.

Implementation. An independent Federal Agency, the U.S. Access Board, has the responsibility to create guidelines that serve as standards for the implementation of the legislation. In 2004, the Access Board announced the new set of revised Americans with Disabilities—Architectural Barriers Act Accessibility Guidelines to meet the requirements of both the old and new legislation. These standards comprise ten chapters of regulations and their coverage includes accessible routes (Chapter 4), communication elements and features (Chapter 7) and recreation facilities and play areas (Chapter 10).

Enforcement. In the case of non-compliance, the matter is directed to the U.S. Department of Justice and pursued through the courts, if this is necessary. However, the US Department of Transportation is responsible for overseeing enforcement in the area of public transport.

A number of countries, including Australia, Canada and many of the European Union countries, have legislation on the rights of disabled people. However, the details of the legislation and the mechanisms for enforcing it vary from country to country.

10.1.2 Accessible Environments: An Overview

The volume of guidelines available from any of the national regulatory bodies indicates the range of issues and complexity of environmental accessibility. The focus of this chapter is accessibility for blind and visually impaired people. However, good practice should apply design for all or universal design principles wherever possible. This involves consideration of accessibility for the whole community, independent of factors such as disability, size, gender, age or ethnic origin. In many cases accessibility features for one group of disabled people have benefits for the whole community, whereas in others different groups of disabled people have differing, possibly mutually contradictory needs. Therefore, it is important to take into account the whole community of potential users when attempting to design accessible environments. For instance, tactile pavements, vibrating warning pads and visual and vibrating indicators of the time remaining to cross at traffic crossings can make these crossings safer for most community members, including children, elderly people, deafblind people, people with cognitive impairments, blind and visually impaired people and non-disabled people. However, while useful in improving safety for most community members, audible warnings and lights can cause stress and reduce safety for people with noise- and/or light-sensitivity.

There are two main components to environmental accessibility for the visually impaired and blind community (see Figure 10.1):

- Good design of the physical environment, which includes the positioning of street furniture to avoid potential hazards, good lighting, the use of colour contrasts and making passageways sufficiently wide for a blind person with a guide dog or human guide.
- The provision of navigation and information systems, for example infrared Talking Signs or a beacon system of the type used in Prague in the Czech Republic.

An important prerequisite for using good design principles for accessible environments is that designers, planners and architects are educated in the principles of accessible environments and *design for all* as part of good design practice. One example of a creative approach to teaching design for all was the Universal Design Education Project (Welch 1995), a funded design education project in departments of architecture, land architecture and industrial and interior design in 22 institutes and universities in the USA. This project aimed to make the principles of universal design an integral component of the student's design education and give them an understanding of the different ways in which a broad range of people use and experience products and the built environment. This project differed from previous work on integrating user needs into design education in its focus on values rather than skills and specific topics.

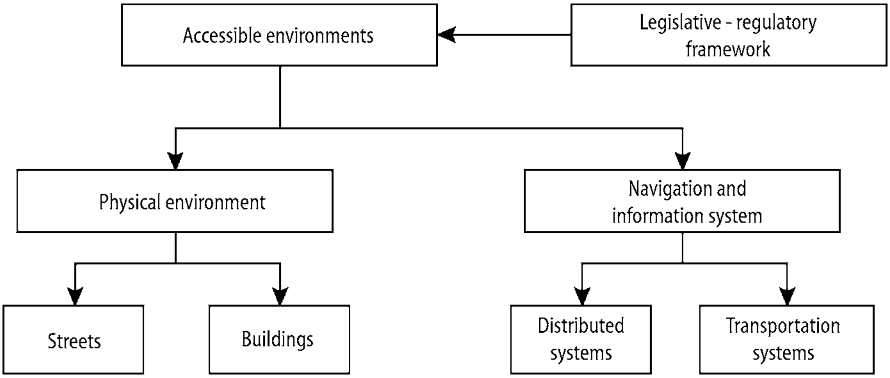


Figure 10.1. Accessible environments: an overview

One important focus of most of the participating institutions was education about the different characteristics, desires and requirements of end-users, particularly disabled end-users. To achieve the necessary input, a common approach was the use of disabled user consultants in the classroom and the design studio. These consultants gave students the opportunity to perceive a product or environment from a very different perspective, and learn about the real needs of disabled people and how following guidelines alone was unlikely to be sufficient to meet these needs. Students found consultants who were able to talk about the details of their lives particularly helpful. Some student stereotypes and preconceptions may have been overcome by presenting the consultants as experts rather than users with unmet needs (Welch 1995).

Another commonly used technique was simulation exercises in which students, for instance, negotiated the campus in a wheelchair or wearing a blindfold. This approach has been criticised by organisations of disabled people and the project organisers (Welch 1995) as trivialising disability issues or reinforcing negative stereotypes about disability. However, such exercises can have a limited role in making students aware of environmental barriers, though not of the experience of being disabled. A preferred approach for learning about environmental barriers, as well as the strategies used by disabled people to overcome them, is for a student to accompany a user consultant around a particular environment and then discuss the experience with the consultant.

The categorisation on which the subsequent discussion in this chapter is based is shown in Figure 10.1.

10.1.3 Principles for the Design of Accessible Environments

While some factors are specific to particular environments, there are a number of general principles that hold for the various different types of environments. These principles include the following, but it should be noted that the list is by no means exhaustive:

- Good lighting—all areas should be well illuminated, generally using diffuse rather than direct lighting and taking care to avoid shadows. As far as possible, a choice of lighting types and the facility to regulate lighting levels should be provided in all rooms in public buildings.
- Well-designed signage systems, which provide both directional information, for example, the direction to the nearest station and location information, for example, each room should have a sign on its door. Whenever appropriate the information should be given in tactile, visual and audio form.
- Tactile paving should be used to warn of hazards and to direct people to facilities.
- The use of colour contrasts, including colour contrasting strips, to provide information and make it easier for blind and visually impaired people to distinguish, for instance, door handles and furniture.
- The provision of wide pathways which can accommodate wheelchair users and blind people with guide dogs, a sighted guide or guide-communicator in the case of deafblind people.
- Good layout design so that facilities, such as benches, litter bins and furniture, do not become potentially hazardous obstacles.
- Matt surfaces to reduce glare.
- Regular maintenance programmes which should be used to reduce the number of avoidable hazards resulting from a poor state of building and streetscape repair.
- Consultation with disabled people, including blind and visually impaired people.

10.1.4 Relationship Between Environmental Information and Navigation Systems and Global Positioning Systems (GPS) Orientation Systems

An integrated environmental information and navigation system is able to provide some of the same types of functions as the widely available commercial orientation systems based on the global positioning systems (GPS) technology. These functions include the location of a wide range of different buildings, environmental features and facilities, and how to get to them. The environmental information system could also provide additional information about these locations, for instance by pressing a button located on the system. This is analogous to the *point of interest* facility of GPS orientation systems for blind and visually impaired people. Currently, unlike orientation systems, environmental information and navigation systems are not able to provide users with routes to a particular destination that they can peruse in advance of travel or direct users to a destination from greater than a relatively short distance, determined by the transmission range of the technology used. They are also not able to search a database for a particular facility, such as the nearest public library, and direct the user to it or to give the user's location relative to appropriate landmarks. There could be benefits in developing integrated information and navigation systems that could provide many of these facilities. One system architecture could have users each using a small receiver with a headphone of standardised design.

10.2 Physical Environments: The Streetscape

Street layout and the position of building and public transport facilities have developed historically with different degrees of planning and regulation in different places. However, it is also only relatively recently that concerted attention has been given to the planning and design of urban environments for accessibility. It should also be noted that some of the potential hazards to blind and visually impaired people, such as large numbers of fast moving vehicles and crowded streets are relatively recent, though they predate concerns about accessibility.

Therefore, existing street layouts and building design put constraints on what can be achieved concerning accessibility, though there are moves to try to ameliorate previous inaccessible design decisions. Concerns for maintaining the character of an area and the practice of 'listing' some buildings as being of particular historical interest with consequent limitations to changes that can be made to them may also put constraints on what can be achieved with regards to accessibility. It is to be hoped that all future streetscape developments will incorporate accessibility considerations.

Another factor that has a significant impact on accessibility is the diversity of vehicles and people using a streetscape. The vehicle types include bicycles, individual mobility carriers, mopeds, cars and trucks, whilst the diversity of human users includes pedestrians, wheelchair users, blind people using mobility devices, people (both disabled and non-disabled) accompanied by children and people pushing children in pushchairs. A particular problem is the fact that in many countries cars and other motorised vehicles sometimes park on the pavement. This can cause a hazard for blind and visually impaired people and can also have the effect of blocking or partially blocking the pavement, particularly for wheelchair users. In areas where there is a deficiency of cycle lanes, cyclists may also cycle on pavements and this could also cause a hazard for blind and visually impaired people. In some countries there are shared paths for cyclists and pedestrians that do not have any delineation between the part of the path for cyclists and that for pedestrians. This could be a hazard for blind and visually impaired people, as well as being generally unpopular with pedestrians and cycling groups.

10.2.1 Pavements and Pathways

Pavements and pathways allow people to move around towns and access facilities, such as shops, doctors, dentists and health centres, public transport, housing and leisure and recreation facilities. Good layout should include features such as ramps instead of steps, lowered curbs and direct routes between facilities with an avoidance of road crossings.

The following factors are important for the design and maintenance of pavements and pathways:

- The surface materials used in pavements and pathways should not become slippery in wet weather. For example, blister blocks in brass or plastic become slippery hazards when wet.
- The edges of pavements and pathways should be clearly identifiable. Kerbs, clear high contrast painted lines and a change in texture can be used to give a visible and/or tactile delineation.
- Steps should be avoided and ramps used.
- Handrails should be provided, particularly on ramped walkways. They can provide guidance to visually impaired and blind people and support to some physically disabled people.
- Pavements and pathways should be regularly gritted and cleared of snow in winter. Currently in some countries, this does not happen, resulting in hazards for all pedestrians, whether or not they are disabled, but particularly for disabled and elderly people.

Tactile pavement blocks

Tactile pavement blocks are commonly used to warn of a hazard or give information about the pavement direction or purpose. There are three main functional types to indicate hazards, give directional guidance and indicate the separate pedestrian and cycle paths on a shared route, as shown in Figure 10.2.

"Corduroy" and blister blocks are used to indicate a hazard. Corduroy blocks, which are marked by a pattern of rib-shaped bars, are used to give warning of a hazard, for example, the beginning of a flight of stairs. Blister paving, where the slabs are covered with a regular pattern of small semi-spherical knobs, are used to indicate the position of a pedestrian crossing and that the pavement has a dropped kerb at the crossing.

Blocks embossed with a rounded bar pattern can be used to give directional guidance. For instance, they could be used to guide visually impaired and blind

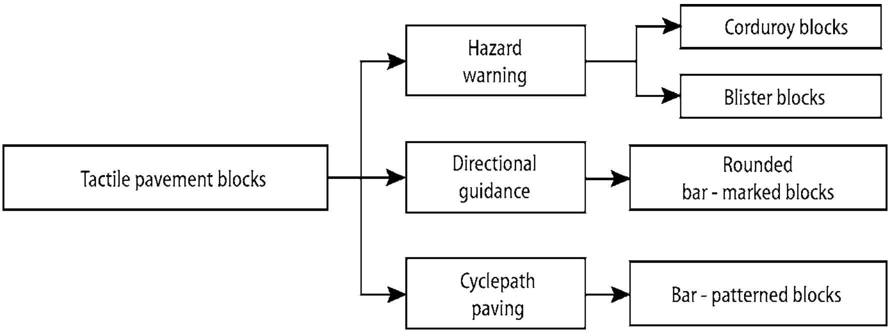


Figure 10.2. The different types of tactile pavement blocks and their applications

people on a safe path across a city square or lead them into and through a pedestrian precinct.

Bar-embossed blocks are used to indicate a shared pedestrian/cycle route. The bars are laid in a transverse direction to indicate the pedestrian path and parallel to the direction of travel to indicate the cyclepath route.

Note that many other groups of disabled people, as well as many non-disabled people can also benefit from both good design and good route maintenance.

10.2.2 Road Crossings

Where pedestrians are crossing a side road and moving in the direction of the main road, it should be recognised that they have right of way. Therefore, vehicles turning from the main road into the side road or the side road into the main road should be required to actively search for pedestrians waiting to cross and then stop, and wait until the pedestrians have crossed the road. Unfortunately, this does not happen currently and existing traffic behaviour privileges motorised vehicles over pedestrians and therefore may discriminate against blind and visually impaired people.

On road crossings, other than at the junctions of very quiet roads, assistive technology systems are often required to facilitate safe crossing of the road. Pedestrian crossing points are frequently indicated by belisha beacons and black and white stripped markings across the road, which are generally called *zebra* crossings. There should also be a lowered curb for wheelchair users, users of mobility scooters and other mobility aids and people pushing children in baby buggies. Many pedestrian crossing points also have traffic lights. These traffic lights may be integrated into the traffic control system at a road junction or solely used to regulate pedestrian or traffic flow. Pedestrians are generally able to instruct the system that they are waiting to cross, most commonly by pressing a button. However, the response time is often very slow and the time allowed for crossing may be insufficient.

Controlled crossings with traffic lights generally have visual indicators, which are normally based on large colour-coded illuminated signs. However, the icons used in different countries differ and include both green stick figures and a large-letter illuminated 'WALK' sign (in the appropriate language) to indicate when pedestrians can cross. These visual indicators are increasingly being supplemented by audible sounds and, less frequently by tactile indicators. In some countries there is also a visual indication of how long remains to cross or to wait to cross. However, crossings without traffic lights generally do not have indicators.

There are also issues of how blind and visually impaired people find crossings. It would therefore be useful for crossings to be equipped with either an infrared Talking Signs type system or a radio frequency identification (RFID) tag which would allow them to broadcast a signal over a moderate distance to be received by a standardised receiver and transmitted to the user either as an audio signal over a headphone or by vibration. The layout of a crossing point, as well as the immediate approach to it should be indicated by tactile pavement blocks of the blister type.

Since the nearest crossing may require pedestrians to go out of their way and pedestrians move considerably more slowly than cars and other motorised vehicles, there would be great benefits to all pedestrians in the development of a device which would indicate to drivers that a pedestrian was waiting to cross and require them to stop. Until this happens, some discrimination against blind and visually impaired people will occur, since they generally require a controlled crossing on even moderately busy roads, whereas sighted people are often able to cross safely at other points. In addition, it should be noted that controlled pedestrian crossings are less frequent than they should be and many roads, which are difficult or unsafe to cross, do not have a controlled crossing.

Crossovers of pavements by motorised vehicles occur where vehicles desire access to car parks, a vehicle entrance to a commercial building, a house driveway or a loading bay. Unfortunately, car and other vehicle drivers fail to recognise that pedestrians should have right of way in these circumstances and often expect pedestrians to stop for them rather than waiting for pedestrians to cross. This situation can therefore create an avoidable hazard for blind people. In some cases, points where traffic crosses the pavement are indicated by a tactile paved area. Tactile pavement indications and colour coding should always be used to indicate the presence of a possible hazard to visually impaired and blind pedestrians.

10.2.3 Bollards and Street Furniture

Bollards and other street furniture can cause a hazard, particularly to blind and visually impaired people, if not appropriately designed and sited. Bollards are used to separate traffic (often parked) from the pedestrian pathway. However, as discussed above, in many countries motorists park on the pavement and will try to circumvent bollards. If too closely spaced or too large, bollards can also prove a hazard to visually impaired and blind pedestrians and an obstacle for wheelchair users, people pushing children and cyclists trying to leave a shared path. The inappropriate positioning or design of bollards can cause a serious hazard to blind and visually impaired people that may result in painful injuries if a collision occurs. Good design (see Figure 10.3) can avoid many of these problems and should include the following:

- Minimum bollard height of 1 m.
- A bollard colour which stands out from the background.
- A band round the neck of the bollard in a contrasting colour to it.

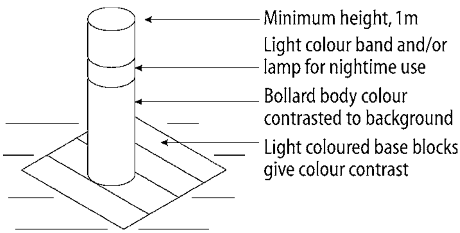


Figure 10.3. Features of good traffic bollard design

- A lamp to illuminate the position of the bollard at night. The illumination should be diffuse and well shaded to avoid causing problems to people with light-sensitivity.
- The Bollards should be physically separated from each other and not linked by chains or ropes.
- An appropriate spacing between bollards to obstruct cars, but not wheelchairs or baby buggies.

Street furniture is the collective term given to a variety of items found on pavements and sometimes on the road, including seating, litterbins, signposts and notice boards. All of these items are potential hazards to visually impaired and blind pedestrians. Some of the potential and avoidable hazards are illustrated in Figure 10.4. This problem can be minimised by good design, including the following:

- Good colour contrast with the dominant background streetscape colours to make street furniture more visually apparent. Bands of contrasting colour on the support poles of notice boards further aid in drawing attention to them. Colour contrasts that are visually pleasing rather than clashing should be chosen.
- Sharp edges and protrusions that can cause injuries should be avoided in the design of street furniture.
- Appropriate location and positioning, so that street furniture does not block the path and can easily be found by blind and visually impaired people who want to use it, for instance, to throw away rubbish or sit on a bench. For example, notice boards on one or two posts should have the posts at the very edge of the path to

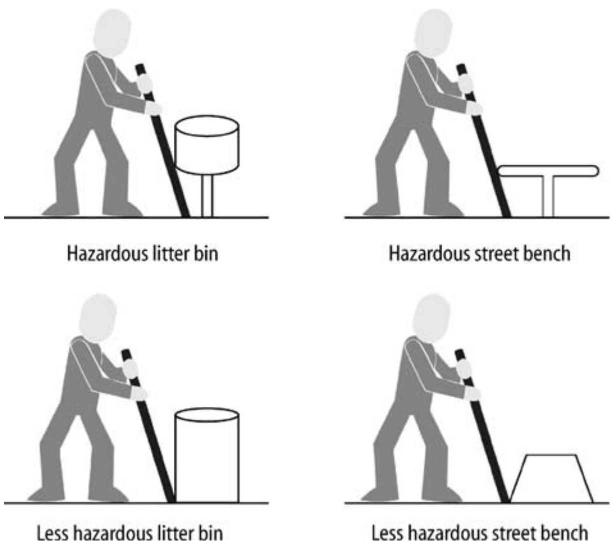


Figure 10.4. Design to minimize the hazard of some street furniture items to the long cane user

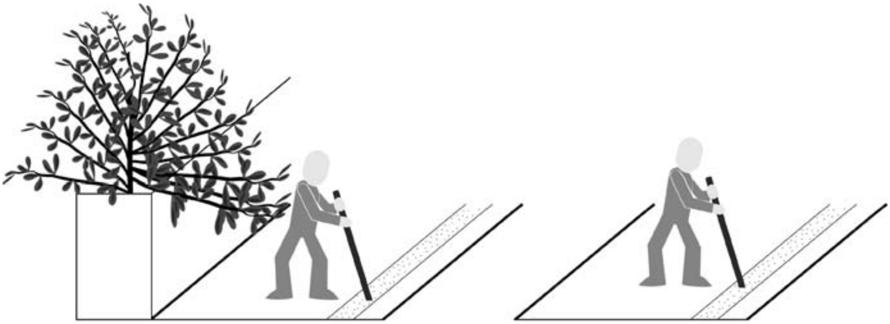


Figure 10.5. Open space way-finding

avoid reducing its width. However, the posts could still interfere with the canes of long cane users.

- Temporary signs of interest to vehicle drivers and cyclists, such as the presence of diversions, should be located at the side of the road and not on the pavement.
- As far as possible, street furniture should have appropriate dimensions to make it more obvious without blocking the path. In practice, the size of most street furniture is determined by its applications.
- The avoidance of potential hazards that can easily snare long canes, such as litterbins supported on a column. The use of a larger bin that sits on the pavement is preferable.
- Consultation with disabled people, including blind and visually impaired people.

Visually impaired and blind people may find open spaces, such as shopping precincts, pedestrian areas and small parks, particularly inaccessible due to the lack of landmarks for way-finding and the hazards posed by the often seemingly random arrangement of street furniture. The provision of well-defined paths marked by tactile paving or other surfaces and high contrast colour strips can make way-finding much easier. Trees and bushes add to the attractiveness of open spaces, but they need to be carefully sited so that they do not obstruct pathways (see Figure 10.5).

10.3 Physical Environments: Buildings

Many buildings were constructed before there was any awareness of the importance of accessibility. Public buildings are now being modified to improve accessibility, but modifications at a later date are not always as successful as designing in accessibility from the start. Building regulations (for public buildings) in most countries now cover accessibility. There is recognition of the need for a percentage of the housing stock to be wheelchair accessible, but there seems to be less awareness of the need for all housing to be as accessible as possible to give disabled people the widest possible choice of where to live and enable them to visit other people freely. Some of the main accessibility principles in the building regulations are discussed in this section.

10.3.1 General Exterior Issues

Access to buildings from the street involves an approachway and an entrance to the building. Buildings may also have external stairways, for instance to allow emergency exit from the building.

Approachways

The approach to the main and preferably all entrances to a building should be wide enough for a wheelchair or two people to walk abreast. The main and preferably all entrances should have an approach which is free of steps, with a ramp with slope of less than 1 in 20 used to overcome any difference in height between the entrance and the road and any gradient. In practice, this may mean that the main entrance has both a series of steps leading directly up to it and a ramp along the front or side of the building that may zigzag to give sufficient length to cover the increase in height without an excessive slope. High-contrast colour edging strips, tactile surfaces and handrails, on both steps and ramps, should all be used to lead pedestrians to the building entrance safely.

Entrances

The entrance to a building should be clearly identifiable and free of obstructions. If the entrance has a canopy, its supporting columns should be designed and sited to avoid creating a hazard. Lighting can also be used to make the entrance distinctive and accessible at night, but it should be diffuse and/or shaded to avoid causing problems to people with light-sensitivity or visual processing impairments. Automatic sliding doors are accessible to most people, but space considerations or other factors may make their installation physically impossible. Indeed, where space considerations permit, a choice of entrance should be provided. Tactile tiling should be used to alert people to the door and lead them to it. Some of the issues to be considered in entrance design include the following:

- The door width should be at least 1.2 m to enable a blind person with a sighted guide or guide dog or a wheelchair user to enter easily. Pillars or other obstacles should not be allowed to compromise this width requirement.
- Revolving doors should be avoided if possible.
- Thresholds to the doorway should be flush with the floor.
- Tactile flooring on both sides should be used to identify the location of doorways.
- Glazed or plate-glass doors should use patterns and high contrast signs to indicate the location of the door.
- Door furniture should be sensibly positioned, be of simple design and have a high contrast colour.

- Spring-loaded doors should have a carefully regulated spring-pressure to enable them to open easily and to prevent them springing back and injuring someone passing through slowly.

Exterior stairways

These stairways are generally provided as an emergency exit (or entrance) rather than for regular use. Therefore, their construction tends to be basic but robust. Despite their infrequent use, it is important that they are as accessible as possible to all members of the community, though they unfortunately cannot be made accessible to wheelchair users. Features that can be used to increase their accessibility include the following:

- Tactile warning surfaces to indicate the start and end of the stair well.
- A clear indication of the edge of the stairs. This should preferably have both tactile and visual indicators, such as a high contrast, tactile edging strip.
- A high contrast coloured handrail on each side of the stairway.
- Adequate illumination of the stairway after dark.
- Appropriate design at ground level to avoid causing a hazard to pedestrians. This should include high contrast signage and a suitable railing or fence around it, if required. The location may make it unnecessary to fence off the stairway.

Signage

All public buildings should have signage in large high contrast lettering, as well as tactile signage. The numbers of private houses should also be larger and clearer than at present. A standard location for displaying them would also be helpful.

10.3.2 General Interior Issues

There are a number of general points that can be considered in order to create an accessible interior environment.

Logical layout

Public (and private) buildings should have a logical layout that facilitates navigation and finding the main facilities. This includes a logical room numbering system, for instance with one number indicating the floor or level and a second number the particular room on the floor. Consecutive numbers should be used for adjoining rooms and the number system should avoid missing out numbers or jumping from high to low numbers and then back to high numbers. Good layout design requires consideration of both people working in the building and people visiting it to access goods and services. In older buildings, it will be possible to modify the existing layout to a certain, though sometimes limited extent in order to improve accessibility. However, even where the layout is less than optimum, lighting, signage and décor can be used to improve accessibility and make the building more attractive and pleasanter to use.

Spatial dimensions for easy passage

A key objective in public buildings is to achieve good circulation areas with dimensions that allow easy passage for all. Where possible, all passages, doors, stairs and lifts should be 1.2 m wide. The following list indicates the minimum passage width to allow passage of particular groups of people:

- Person with walking stick 750 mm.
- Person with a guide dog 1100 mm.
- Person with a sighted escort 1200 mm.
- Person using two crutches 900 mm.
- Wheelchair user 900–1000 mm.

Hazard-free passage

This has two components: the building structure, which cannot be modified without structural changes and moveable furniture and obstacles. Structural obstacles include support columns, fixed door widths, steps, and stairways. The dimensions and structural requirements of the building, as well as cost considerations, may make it impossible to change these features. However, the use of good colour contrasts, panelling, tactile paving and appropriate lighting can draw attention to these features and minimise their hazard potential. If appropriately designed, these features can also serve as orientation points. Moveable obstacles include both semi-permanent features, such as seating, (low) tables, drinking water fountains, fire extinguishers and sand buckets and temporary obstacles left by building users. All these features need to be appropriately located away from passageways so that they avoid causing a hazard and can easily be located and used by both blind and visually impaired people and wheelchair users. All building users should be careful not to leave equipment, clothing or anything else where it can cause an obstacle. Accessibility also requires the maintenance of the building in a good state of repair to avoid, for instance, the need to use buckets to catch drips. Particular care is required when repairs take place inside a building that is in use to reduce potential hazards, for instance, due to the presence of builders' ladders, and to inform building users of the need for particular care.

Lighting and décor

Lighting and décor, including the use of good colour contrasts, can make buildings more accessible to blind and visually impaired people. Tactile floor and wall markings and surfaces are also important for improving orientation and ensuring a safe passage round the building for blind people. There seems to be an increasing tendency to use very bright direct lighting in the entrances and passageways of some public buildings. This use of direct lighting should be avoided, as it makes them inaccessible to people who are light-sensitive or have some visual processing impairments. Lighting and décor also have an important role in contributing to the attractiveness and welcoming atmosphere of public buildings. This should be considered an important component of accessibility rather than a competing or conflicting requirement, since public buildings are intended to be used and a building that feels unwelcoming and unattractive may turn away some potential users. Lighting and décor are discussed in more detail in Section 10.3.3.

Signs

Finding a route through a building is often a matter of following signs. Therefore, a well-designed signage system that provides information in a variety of formats and has a logical structure is required. Otherwise, all users, especially blind and visually impaired people, are likely to get lost and have a frustrating time trying to reach their destinations. Signage is particularly important in large buildings and complexes of buildings, such as university campuses. Signage can be divided into two main categories: directional and locational. Both types are equally important, since users require signage both to find the lecture hall they are looking for and to identify it so they know they have reached the correct location. All signs should be appropriately located to make them perceptible and should not be obscured. Where possible, combination of tactile signs, large high contrast lettering, symbol signs, picture or icon signs and audio signs (to be received over a headphone of standard design) should be used.

Building services

Building services is the usual term given to all the additional service items needed to run a building; items like power systems, water supply and lifts come in this category. Often these services are the responsibility of a building services engineer. Two factors need to be considered concerning the accessibility of some building services, such as security door systems, lifts, escalators and moving pavements (travelators): all building users, including blind and visually impaired people, should be able to locate and use them easily and potential hazards should be avoided. This requires the use of signage, tactile paving and appropriate location.

Special purpose rooms

Many of the rooms in public buildings serve a designated purpose, such as a canteen or a dining room, office, hotel bedroom, interview room or toilet. In addition to general accessibility considerations, particular attention should be given to ensuring that all functions of these rooms are fully accessible. It can be seen that the list of accessibility guidelines easily becomes extensive when special purpose rooms are considered and, for this reason, no further discussion of these special purpose room issues is given in this chapter.

Soundproofing

Unfortunately, accessibility guidelines, such as those resulting from the Americans with Disabilities Act (ADAAG 2002) generally do not consider the issue of soundproofing. However, it is covered in many building codes. Appropriate soundproofing both between rooms and on external walls and windows to cut out external sounds is important for both disabled and non-disabled people. As well as reducing stress and physical and psychological illness from external traffic and other sounds and the potential for disputes between neighbours, a quiet environment makes it easier for hearing impaired people to hear signals of interest, such as conversations and visually impaired people to perceive relevant auditory environmental cues.

The weighted sound reduction index gives a measure of the extent to which, for instance a wall, insulates sound by attenuating sound travel through it. It is determined over a frequency range of 100 Hz to 3.15 kHz (NI, undated; AZBUILD, undated), which covers the main speech frequencies, but does not cover the higher frequencies perceived by the human ear. Typical values in building codes are 55 dB for internal walls.

10.3.3 Lighting and Décor

Lighting

Good lighting is very important for illuminating the interior of a building, particularly at night. It can also be used to improve accessibility by indicating or drawing attention to the location of particular facilities. Lighting and décor also contribute to the ambience and attractiveness of a building. General points to be considered in designing the lighting system include the following:

- Dramatic changes of lighting levels from one public area to the next should be avoided. Uniformly distributed lighting facilitates the passage of visually impaired and blind people around the building. People with certain types of visual impairments may adapt slowly to changes in light levels and this can create a hazard if lighting levels change suddenly.
- Glare can be caused by light reflection from polished surfaces such as floor coverings, windows, computer screens and furniture. This can lead to temporary blindness for people moving through these areas and therefore care should be taken to avoid possible sources of glare.
- Lights should be adequately shaded. Very bright lights with little shading can cause problems for people who are light-sensitive or have certain types of visual processing impairments and should therefore be avoided. Unfortunately, this type of lighting seems to be used increasingly in entrance halls, access corridors and lifts.
- As far as possible, lighting systems should be designed to avoid shadows or dark areas caused by objects in the path of an illuminated source, as they can obscure pathways and create visual illusions, leading to confusion and possible accidents. This is particularly important in public access areas and should be checked at different times of day and night.
- Different types of lighting, including overhead lighting and task lighting, should be provided to enable users to choose the best type or combination of lighting

Table 10.1. CIBSE recommended lighting parameters

Location	Task/function	Maintained illuminance (lux)	Limiting glare rate (UGR)	Minimum colour rendering (R_a)
Entrances	Halls, lobbies, waiting areas	200	22	80
Enquiry desks	Face-to-face transactions	300	22	80
Airport: escalators and travelators	Local pedestrian movement	150	22	80
Kitchens	Food preparation	500	22	80
Offices	General clerical work	500	19	80
Public car parks (indoor)	Parking areas	75	–	20
Railway stations	Covered platforms	50	28	40
	Open platforms	10	–	–
	Waiting rooms	200	22	80
Theatres, Concert halls	Foyers	200	–	–
	Auditoria	100	–	–
	Toilets	200	25	80

to meet their needs. Many visually impaired people can find task lighting, which focuses on the task being performed, very useful, whereas some light-sensitive people and people with visual processing impairments may not be able to use task lighting.

- The ability to adjust the level of lighting, particularly in special purpose rooms, such as offices, is very useful to enable all building users to have appropriate levels of lighting. This would be very useful in hotel bedrooms, where the lighting is often very dim.
- Building regulations and guidelines generally include recommended lighting levels for specific building areas. They should be consulted and provide a useful indication. However, the recommended lighting levels do not always provide sufficient illumination to meet the needs of visually impaired people. A typical set of guidelines for recommended lighting values is that produced by the U.K. Chartered Institution of Building Services Engineers (CIBSE), a section of which is shown in Table 10.1 (CIBSE 2002).

Illuminance

The science and measurement of lighting is a well-developed aspect of building services engineering. To present detailed information about this field is beyond the scope of this book but it is very useful to understand the main technical parameters involved. Definitions of the full range of parameters used in lighting can be found in a joint publication of the Commission Internationale de l'Eclairage and International Electrotechnical Commission (CIE 1987). The first parameter is illuminance. This quantifies the amount of luminous light flux received at a surface and is usually measured in lumens per square metre or lux. Thus, the lighting levels in a room or interior can be correlated to the supplied illuminance.

Units for light

The measurement units for light are based on the solid spherical geometry of a point source radiating light uniformly outwards to illuminate a concentric spherical surface, as shown in Figure 10.6. The sphere enclosed by this surface has a solid angle of, 4π steraradians (st).

The light source emits radiation. This gives a radiant flux across the sphere, which is the radiation energy transported per unit time and therefore measured in watts. This energy flux produces luminous flux on the surface, which is measured in lumens. The luminous intensity is the luminous flux per unit solid angle and this is measured in candela, with one candela equal to one lumen per steradium. Alternatively, a candela can be defined as the luminous intensity in a particular direction of a single frequency light source of 540×10^{-12} Hz with a power of 18.3988 mW energy intensity in that direction of 1/683 W/sr. The total flux in lumens can then be obtained from the luminous intensity in candela by multiplying it by the solid angle in steradians into which the light is emitted. The illumination or brightness is the surface density of the luminous flux received. It is measured in lux with one lux equal to one lumen per square metre.

Most lighting designs are described in terms of illuminance quantities, namely the amount of luminous flux (lumens) per unit area reaching the various surfaces in an environment. However, the visual system also responds to qualitative factors such as glare and colour in determining whether the lighting promotes visual comfort and visual satisfaction. These two issues provide the next two parameters for discussion; the limiting glare rating and the index of minimum colour rendering.

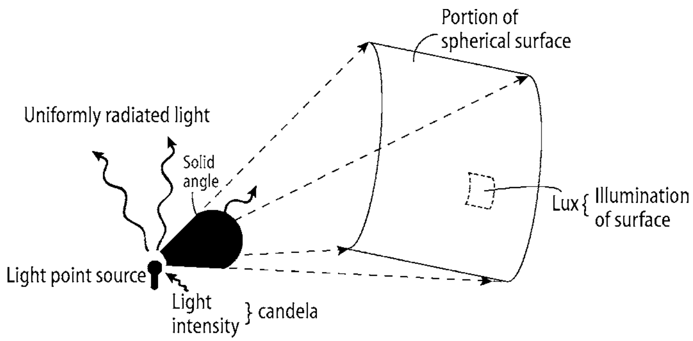


Figure 10.6. Units for light measurement

Glare

Glare results when there are bright and dark areas in the field of view or some areas in an environment that are much brighter than the general level of brightness. Glare can be divided into the following three categories:

- Blinding glare, for instance from staring at the sun. This results in temporary total loss of vision and leads to temporary visual impairment. It is generally not relevant in indoor environments.
- Disability glare, which is the presence of excessive illumination that reduces vision to the extent that objects and details can no longer be seen. It could be caused by oncoming car lights.
- Discomfort glare, which results when an interior experiences a much higher level of illumination than is necessary for its role. Although not dangerous in itself, it is very irritating and can lead to fatigue if experienced over an extended period. A common cause is prolonged exposure to an over-illuminated interior, which results in a general feeling of discomfort from being in the area; hence the name of discomfort glare.

Units for glare

Based on experimental research, discomfort glare has been shown to be dependent on the luminance and size of the source of illumination, as well as its position in the field of view and the general level of luminance in the (interior) environment. These factors have been combined into a formula from which the degree of discomfort glare can be computed as a unified glare rating (UGR). Thus for a given illuminated interior, it is possible to compute its UGR value and compare this with a recommended limiting UGR value. If the recommended limiting UGR value is exceeded, then the lighting system should be redesigned. In general disability glare will not be a problem if discomfort glare is maintained within appropriate limits. In addition to other sources of glare, many modern appliances, such as mobile phones and MP3 players, are backlit and the resulting glare can be very distressing to people who are light-sensitive. Limiting glare rating values for disability and discomfort glare for a variety of appliances are given in CIBSE (2002).

Colour

An important feature of surface colour is its reflectance. Colour systems are used to describe and specify the effects of a particular colour scheme. In the Mansell colour system, three quantities are used to specify colour: value, chroma and hue. Colour value relates to its lightness and hence to its reflectance. Colour chroma defines colour strength and colour hue specifies the type of colour, for instance whether it is red, yellow, green, blue or purple. These terms can be used to specify and describe the appearance of the coloured surfaces of a room.

The following two physical properties: (i) the apparent colour of the emitted light and (ii) the colour rendering are used to quantify the colour properties of light. These properties are defined by the Commission Internationale de l'Eclairage

(CIE) (CIE 1995). In specifying the properties of interior lighting systems, it is the index of colour rendering that is generally most important.

Units for colour rendering

Colour rendering is the ability of a light source to render or replicate the colours of surfaces. The CIE general colour-rendering index is a measure of the accuracy with which a light source or lamp is able to reproduce a set of test colours when compared to the test colour reproduction of an appropriate standard light source. The colour rendering index, R_a ranges from 0 to 100, with 100 signifying perfect agreement in colour rendering. Only the highest values indicate good colour performance, with a colour rendering index of less than 80 signifying that a lamp should not be used in areas where people are working for extended periods.

The *Code for lighting* document (CIBSE 2002) provides lighting specifications for a very wide range of interior and exterior light systems. The general reader might be surprised to find lighting specifications for locations as diverse as bakeries, autopsy rooms, steel rolling mills and cold stores. The specifications are given in terms of (i) the illuminance in lux, (ii) the limiting glare rate on the UGR scale and (iii) the minimum colour-rendering index, R_a. Table 10.1 shows typical recommended lighting specifications.

Décor

Décor is very important. It both contributes to the attractiveness and ambiance of a building, room or facility and provides clues to support way-finding and safe passage. These two functions should be seen as complementary rather than in competition, though compromises will sometimes have to be made. However, it needs to be recognised that accessible buildings should also be attractive ones. Strong colour contrasts are required to assist visually impaired and blind people in differentiating building features, such as walls from floors, and locating doors, door handles and light switches. These colour contrasts should be chosen to be harmonious and attractive, for instance by using a very dark shade of a particular colour against a very pale shade of the same colour.

Figure 10.7 illustrates how poor décor selection could lead to confusion and a potential accident point in a public library. The carpet flooring and the built-in textile-covered reading seat between two fixed bookcases are an identical carmine red shade and of a similar texture. As a result, visually impaired people could miss seeing the seat due to the lack of colour contrast and try to reach to the open space beyond through the 'gap' between the bookcases, leading to minor accidents.

Mirrors and glass panels are sometimes used to give the impression of more space. Since visually impaired people can find this very disorientating and may try to walk through the apparent open vista in front of them, the use of mirrors and panels in this way should be avoided.

Some simple principles for the interior decoration of public buildings are given below (Barker *et al.* 1995):

- Wall finishes should be matt and use pale tones.

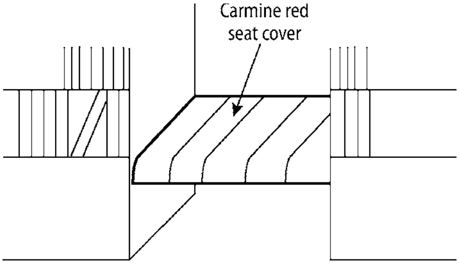


Figure 10.7. Poor décor and design for a public library reading seat

- Floors and floor coverings should be plain and have a colour contrast to the wall colour.
- Uncovered floors should not have a gloss finish to reduce glare.
- To help the uniform distribution of lighting, ceilings should be in pale tones.
- Doors should contrast with the walls and door handles and locks should be colour contrasted with the door colour. These colour contrasts make it easier for visually impaired people to find, open, lock and unlock doors.
- Sockets, switches, handles and similar small items should be colour contrasted to their background so that they can be found easily.
- Chairs, cushions and sofas should be colour contrasted to the colours of the floor and wall, for instance by using a much darker shade of the same colour, to make them easier to locate.
- The use of mirrors and glass panels to give the impression of more space should be avoided. This discrepancy between appearance and reality could be hazardous to visually impaired people, who might try to walk through the panels or mirrors into the apparent space beyond.

10.3.4 Signs and Notices

Signs and notices can be classified in a number of different ways, including the following:

- The sensory modality used, for example, visual, audio, tactile or a combination of these.
- The type of technology used in the sign, such as, infrared, mechanical ('low-tech'), Bluetooth or visible light.

As far as possible, all signage information should be presented in visual, audio and tactile format, with the visual content including both high contrast clear lettering and symbols or pictures. For instance, the sign for a lift should have Braille embossing as well as the word 'LIFT' in large high contrast lettering. In some cases, this will involve the use of separate signs for the different sensory modalities.

This section will consider mechanical low-tech signs, which are generally visual and/or tactile. Audio and 'high-tech' tactile and visual signs will be discussed in the subsequent sections. The visual part of the sign should include symbols or pictures as well as high contrast clear lettering. The message should be short, simple, easily understood and unambiguous. This is always important, but particular care needs to be taken to ensure that there is no ambiguity in any messages with safety connotations.

The following recommendations, with added italicised comments by the authors, resulted from research by the RNIB on legibility (Barker *et al.* 1995):

- White lettering on a dark background generally improves legibility for people with a visual impairment. *However, some people can find that this colour combination causes glare and prefer dark lettering on a white background.*
- The majority of letters should be lower case, as it is generally easier to read than upper case lettering.
- Signs should be positioned at eye level with easy access for close-up viewing. *However, the RNIB suggestion of an appropriate eye level of 1.4–1.6 m is most satisfactory for standard size adults and less so for wheelchair users, small or very tall adults and children.*
- Care should be taken to ensure signs are not obscured by environmental or building features or furnishings.
- Glare from signs should be reduced by using matt surfaces and avoiding glass cases.

Standardisation of the position of commonly used types of signs, such as street names would be helpful. Currently street name signs in some cities or parts of cities are fastened to buildings at adult eye level, whereas in other places they are located on short poles about 0.5 m above the ground. A standardised location would make it easier to know where to look for the street name sign. Signs should also be located close to the facility they are indicating. For instance, lift signs should be located close to the lift call button to enable blind people to access the tactile content of the sign and the lift call button together. Both the sign and the lift call button should be at an appropriate height to be accessed by wheelchair users.

Signs with tactile markings need to be located at an appropriate height to be touched easily, including by blind wheelchair users and small blind adults and children. This will mean that signs with both visual and tactile information should be located at a height of about 0.6–0.8 m and therefore considerably below eye level. To avoid taller adults having to bend down to reach the sign, two sets of tactile signs at different heights would be useful. This is particularly important in the case of Braille, where signs that are too low down will appear upside down. Grade 1 Braille can be used for single words, but grade 2 Braille should be used for longer expressions to reduce the size of the sign. Embossing is easier to feel than engraving and an appropriate tactile height for embossing letters and symbols is 1–1.5 mm.

Table 10.2. Sign colour contrast guidelines

Background	Signboard	Legend
Red brick or dark stone	White	Black, dark green or dark blue
Light brick or light stone	Black/dark	White/yellow
Whitewashed wall	Black/dark	White/yellow
Green vegetation	White	Black, dark green or dark blue

The preferred colour contrasts for typical sign location backgrounds are shown in Table 10.2.

10.3.5 Interior Building Services

Good design is required to ensure that door entry and exit systems, lifts (elevators), escalators, moving pavements (travelators) and emergency call systems can be easily used by visually impaired and blind people and are not a source of potential hazard.

Entry and exit systems

Many modern buildings use mechanical and electronic door release systems to increase building security. There are a number of different means of operating the door release systems, including plastic cards with a magnetic strip, a system of coded push buttons, often in combination with a magnetic strip card, and metal key systems. The systems need to be easy to locate and use. Therefore, the system should be positioned at about 1.0–1.2 m with a tactile strip from a height of about 1.4 m downwards to lead blind and visually impaired people to the keypad and card system. Key pads should have the numerals 1, 2 and 3 on the top row and a small bump on the button for the numeral 5. The keypad buttons should be of a good size with large high colour contrast lettering and the buttons should have a high colour contrast with the door. New systems should be designed so that the card can be inserted in any direction and orientation. When this is not the case, the card orientation should be indicated by visual and tactile markings. Visual and tactile markings may also make it easier to locate the slot for inserting the card and indicating the direction of insertion.

Particular care should be taken with the design of emergency exit doors, the controls of which have become increasingly complex. In many cases, electronic pads and bush buttons are used to work the emergency door release mechanism and an appropriate location and high contrast highlighting are required to make them easier to detect and use. In the case of corridors with handrails, the electronic release mechanism should be located close to the handrail next to the emergency exit door.

Emergency call systems

Emergency call systems are now present in a wide range of building facilities, including lifts, toilets, staff rooms and escalators. As with other facilities, they should have a sensible location, clear identification and ease of use. This requires a high contrast colour to the background, a large push pad area and easy operation with a minimum level of pressure.

Accessible lift (elevator) example

The guidelines presented in this section on décor, lighting and signage will be illustrated by the application of the comprehensive assistive technology (CAT) model of an accessible lift. This is found in Table 10.3 which should be self explanatory.

Escalators and moving pavements (travelators)

Moving pavements (travelators) are horizontal escalators and therefore similar principles of accessible design can be applied to both escalators and travelators. Escalators are used in a wide range of different types of locations, including airport buildings, department stores, shopping centres and multistorey office buildings, whereas moving pavements are mainly used in airports and some railway stations. In some countries, such as the Czech Republic, guide dogs are trained to assist blind people in using escalators, whereas in others, such as the UK they are not. Consequently, blind people who use guide dogs may find escalators difficult or impossible to use. There is always a stationary avenue alongside a travelator to allow people to walk or to travel in the small carts provided by the airport. Therefore visually impaired and blind people and their guides (whether human or canine) can use these alternatives. However, it should be noted that the travelator has the advantage of allowing faster mobility than the stationary avenue and that the alerting sounds and lights of airport carts can make them unuseable by some travellers, as well as increasing the stress experienced by these travellers when they are in the vicinity. Escalators also cause difficulties for wheelchair users. Therefore, as far as possible, alternatives to escalators in the form of lifts and ramps should be provided, though this may be difficult in older buildings due to the lack of space.

Tactile floor markings should be used to direct blind and visually impaired people to escalators and travelators and their start should be indicated by a combination of tactile warning pads and high contrast large lettering notices. Many travelators have localised audio alerts to indicate that their end is approaching, but this is not the case for escalators. In both cases, a tactile indication, such as a very light vibration of the floor, would be useful, though escalator, but not travelator, users experience a physical sensation of the steps levelling out. High-contrast white or yellow strips should be used to mark the pavement edges of travelators and the sides of escalator steps. The escalator and travelator system should be well illuminated and their entry and exit areas should be clear from obstructions, including litterbins and advertisement hoardings. There may be a need for an improved design to ensure a smooth transition from the escalator or travelator to the stationary floor.

Table 10.3. CAT model for accessible lift

Attribute	Component	Factors
Context	Cultural and social context	Assistive technology is available, but information about available products is sparse
	National context	National building guidelines available, but compliance is voluntary
	Local settings	Indoor building environment Noisy environment, confined space
Person	Social aspects	May travel with a human guide, a guide dog or alone, possibly with a long cane or electronic travel aid
	Attitudes	Adventurous and willing to seek help from assistive technology systems
	Characteristics	Visually impaired or blind with no or only very mild hearing impairment Reads using large print, Braille or Moon
Activity	Mobility	Travelling in an unfamiliar lift system
Assistive Technology System	Activity specification	*Task and User specifications* A visually impaired or blind person should be able to call the lift, enter the lift, direct the lift to the desired floor and leave the lift safely on the correct floor In the event of lift failure, a visually impaired or blind person should be able to activate and use the emergency call system without problems
	Design issues	Design approach: considers solutions for visually impaired and blind people rather than a design for all approach Technology selection: visual, audio and/or tactile solutions
	Assistive technology system Environmental interface	Tactile flooring material around entrance to lift Signage in large high contrast letters and tactile symbols Braille embossing on call lift button Safety checks on door closing; audio announcement of doors opening and closing Lift cabin: minimum dimensions 1.4 m deep by 1.1 m wide Safety rail around cabin; lighting of uniform intensity; non-reflective lift cabin walls (not mirrors)
	Assistive technology system Processor components	Onboard lift controls use large push buttons Power and speakers for onboard audio system
	Assistive technology system Human–technology interface	Lift control buttons at 1.0–1.2 m height; tactile numeral symbols, Braille embossed buttons; buttons should have colour contrast for high visibility Floor number and door status (*e.g.* opening or closing) announced by audio system. Visual floor indicator in large colour contrast letters Emergency call button—highly visible and tactile markings on button. Audio messages for action in event of lift failure
	End-user issues	Mode of use: occasional Training requirements: None Documentation: None

10.4 Environmental Information and Navigation Technologies

Successful navigation of both the exterior and interior built environments is largely dependent on good quality signage, which should provide a combination of audio, tactile and visual information. This section considers audio signage, with Section 10.4.1 presenting an overview whilst Section 10.4.2 provides a description of some specific technologies.

10.4.1 Audio Information System: General Issues

Categorisation of audio signs

There are a number of different ways to categorise audio signs, including the following:

- The sensory modalities used to transmit the information; these include:
  1. Single delivery modality, that is, solely audio information.
  2. Dual delivery modality, that is, audio with visual or more rarely tactile information.
  3. Triple delivery modality, that is, audio with both visual and tactile information.
- The type of technology used to transmit the environment information. Currently there are systems that use Bluetooth, wireless local area networks (LANs), infrared and visible light.

Technological aspects

The following principles of good design are common to all audio sign systems independently of the technology used.

Easy location of the sign or notice from a distance The two main options are illustrated in Figure 10.8 and given as:

1. Signs that are constructed as a sound beacon transmitting an audio signal to guide the traveller to its location.
2. Signs that transmit a continuous signal that can be intercepted by a handset. The transmitting medium may increase in strength as the user approaches the sign. An example is the Talking Signs technology based on infrared technology.

Sign activation There are three main cases, depending on whether the sign transmits continuously or requires to be activated and how it is activated. These three cases are illustrated in Figure 10.9 and listed as:

1. Activation by pressing a button located on the sign. This means that the user needs to be sufficiently close to the system to press the button and listen to the message transmitted.

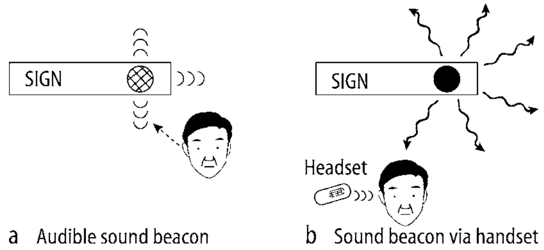


Figure 10.8a,b. Finding an active sign or notice in the environment

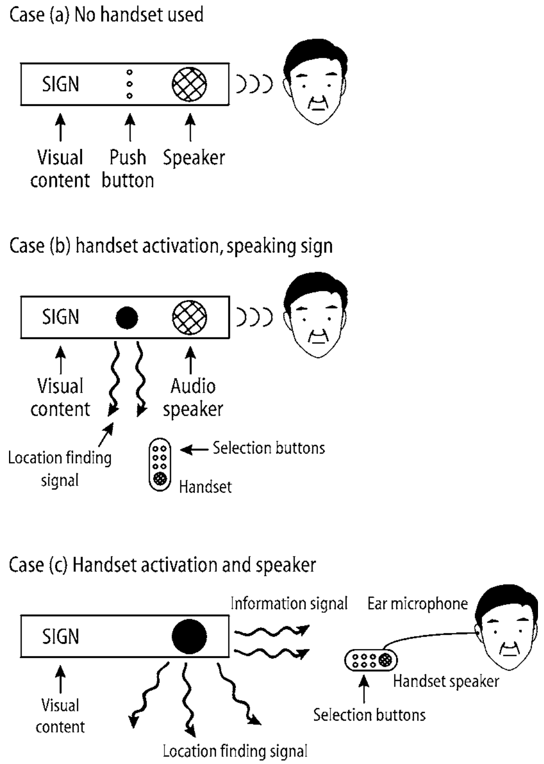


Figure 10.9a–c. Methods for accessing information in an audio sign or notice system

2. The use of a handset to activate the sign, which then delivers its audio message to the user from a speaker embedded in the sign. The user needs to be sufficiently close to the system to activate the sign and listen to the information message transmitted.
3. The use of a handset to intercept the continuously transmitted signal from the sign. The handset then delivers the message to the user. This means the user could be some distance from the sign and still receive the information content

of the sign. The handset could have a small embedded speaker or an earphone jack plug socket to deliver the audio signal.

Sound reception The audio message is heard by the user either on a earphone connected to the handset, or on a speaker, which could be located on either the handset or the sign. The use of a single earphone is generally preferable, as it reduces the impact of interference from environmental background noise and prevents disturbance to other pedestrians. As audio signs become increasingly widespread, avoiding disturbance to other pedestrians is going to become more important. The use of a single earphone also allows the user to access other audio environmental information.

Standardisation There are clear benefits to users in technological standardisation to allow all the different types of audio signs to be received by one handset, which would then process and convert the signal to transmit it to the user in audio or tactile form.

Coherent design of audio signage Currently audio sign systems are sparsely distributed and therefore tend to be installed on a local basis. There is a need for a more systematic approach involving national systems which can then be combined into a global system. This should include consideration of avoiding interference between different audio signs and the complementarity of audio signs and orientation systems, for instance based on global positioning systems.

10.4.2 Some Technologies for Environmental Information Systems

A number of different technologies can be used to transmit information from beacons that are embedded in the physical environment. In order to perform satisfactorily transmission technologies should have a number of properties, including the following:

- An adequate transmission range for the application.
- Non-invasive and/or imperceptible transmission to avoid disturbing other people.
- Avoidance of injury to people, animals and the environment.
- Mature and well understood technology.
- The easy availability of inexpensive system components.

Ultrasonic technology is generally not used in audio signs and therefore the different transmission technologies are based on particular frequency ranges of the electromagnetic spectrum. Table 10.4 shows typical properties and the frequency and wavelength ranges for the four main communications media used.

A detailed description of a public transport information system that uses Bluetooth and mobile telephone technology is presented in Chapter 11. Systems that use infrared and visible light transmission technologies are described in this section.

Table 10.4. Properties of four communications media

Type	Frequency and wavelength ranges	Properties
Bluetooth	2.4–2.484 GHz and 12.5–12.07 cm	Omnidirectional, interior/exterior use, penetrates walls, doors
Wireless LAN	2.4 GHz and 12.5 cm (Wifi/802.11b and HomeRF) 5.15–5.35 GHz and 5.83–5.61 cm (802.11a, WLAN and HiperLAN/2)	Omnidirectional, interior/exterior use, penetrates walls, doors
Infrared	3×10^{12}–430×10^{12} Hz and 0.1 mm–750 nm	Unidirectional, interior/exterior use, blocked by walls, doors
Visible light	4.3×10^{14}–7.5×10^{14} Hz and 750–400 nm	Omnidirectional, interior use only, blocked by walls, doors

10.4.2.1 Infrared-based Environmental Information System

There are several commercially available systems that use infrared technology to create a beacon system to provide information messages to a user's handset. These systems include Talking Signs®, Infravoice, and the Easy Walker system. The Talking Signs system has been installed in a number of different types of locations in several different countries. The discussion of the Talking Signs system in this section is used to illustrate the operational principles of infrared-based systems and their potential flexibility.

Talking Signs

Talking Signs provides a repeating, directionally selective voice message, which originates at the sign and is transmitted by infrared light to a hand-held receiver some distance away. The directional selectivity is a characteristic of the infrared message beam where the intensity and the clarity of the message increases as the sign is pointed at or approached. This ensures that person using the Talking Signs system can obtain feedback about their relative location to their goal as they move towards it. Talking Signs are light and small, easy to install, consume very low power, and are easy to program with human voice or synthesised speech messages. Figure 10.10 shows a Talking Signs system in action. The user is pointing a receiver handset at a wall-mounted beacon to obtain information about the stairway.

The infrared system uses light-emitting diodes to transmit digitally encoded speech messages that are intercepted by a hand-held receiver and then a speaker in the receiver unit relays the message to the user. Each infrared transmitter consists of a rectangular plastic box containing the message unit, transmitter driver electronics and three LEDs. The pre-recorded human speech frequency message modulates a 25 kHz infrared carrier signal. As the infrared beam carrying the speech message is transmitted, it spreads out as a cone from the infrared diode, becoming wider as the beam moves away from the source. Tuning the transmitters and the LED arrays allows control of the maximum distance at which the Talking Signs message is received, the direction of transmission, and the area

Figure 10.10. Talking Signs® in action (photograph reproduced by kind permission of Talking Signs Inc. USA)

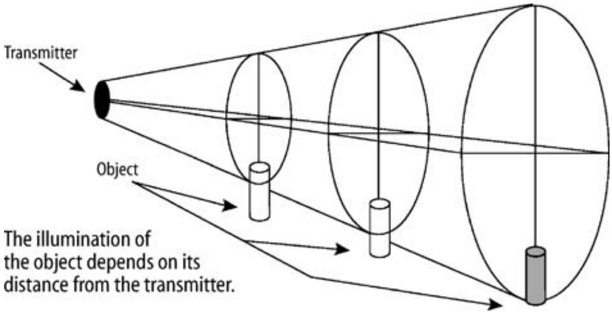


Figure 10.11. Cone of infrared transmission signal

the message covers, as illustrated in Figure 10.11. Because different signs have different functions, the range and dispersion angles of each sign are adjustable.

The hand-held receiver consists of a box containing a frequency modulated (FM) discriminator, an amplifier, an internal speaker and a photo-detector at its front end. The message is detected when the receiver is pointed in the direction of the sign transmitter. A digitally recorded message is heard whenever the sensor aperture on the front of the receiver is pointed in the direction of the infrared transmitter while the receiver is activated. Thus, Talking Signs are a directional system whose messages are received only when the user activates the receiver. Therefore, system users do not receive unwanted information. However, other pedestrians in the vicinity of the user may hear the message, making the use of a single headphone preferable. This would also prevent the information disappearing in environmental noise. The handset has an on/off switch, a volume control and a jack port for headset use. It is powered by a 9-V battery and can be carried on a cord around the neck to free the hands for other tasks. Figure 10.12 shows a Talking Signs transmitter beacon and a handset with internal speaker.

This type of beacon and information access technology has a number of potential and actual applications, including the following:

1. Way-finding and information systems for interior and exterior environments. Exterior environments include shopping centres, railway and bus stations, whereas interior environments include hotels, hospitals and government of-

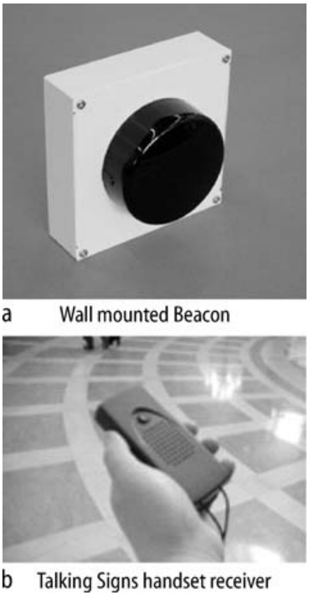


Figure 10.12a,b. Talking Signs® beacon and handset receiver (photograph reproduced by kind permission of Talking Signs Inc. USA)

fices. A centralised system would typically be based on groups of ten transmitters powered through a centralised control unit. Stand-alone units could be used in locations that are not easily linked to a centralised system and for one-off applications. These single units would have a dedicated transformer to provide 12-V power, a recorded card and dedicated transmitting head.

2. Personal way-finding systems, comprising a transmitter with messages recorded by the user and a personal receiving handset. This could be used by individuals to mark and locate personal destinations, such as a hotel room door, a train seat or a restaurant table.
3. The Talking Signs® technology has been incorporated into the electronic destination panels of buses. This enables the receiver handset to be used to identify individual bus services from a distance of around 30 m, thus facilitating independent bus travel. Figure 10.13 illustrates such a bus identifying system.
4. A modified Talking Signs system has been used in pedestrian crossing systems to provide commands for safe crossing and additional information about crossing locations and traffic conditions.

Users would benefit from standardisation to allow the same handset to be used with all beacon and information access applications. Other pedestrians, particularly noise-sensitive ones, would benefit from the internal speakers being replaced by a headphone, especially as Talking Signs and analogous systems become more widespread. It should also be noted that the combination of a small hand-held box with a headphone resembles a mobile phone with earpiece or an ipod and therefore does not draw attention to the user.

10.4.2.2 Optical Technology for an Environmental Information System

The Talking Lights® system uses optical technology to transmit the signal. The light from an ordinary light fitment is used to encode information to create a communication channel. The modulation does not interfere with the production of illumination from the light and the modulation frequency is chosen so that there is no visible light flicker.

The technology has three components:

1. A fixed modulating light system that emits the locator signal
2. A portable receiver to receive the locator and information signals
3. Signal processing software to process the light signal and convert it to audio format

The receiver handset decodes the information from the modulated light signal and presents it to the user. The system can be used to provide a way-finding system around buildings.

10.5 Accessible Public Transport

All types of public transportation systems need to be accessible to all sections of the community. This is particularly important for visually impaired and blind people, who generally do not have bicycles, cars or mopeds and therefore rely on public transport, including buses, trams and rail, for local and national travel. This section discusses the accessibility considerations of public transport for blind and visually impaired people.

There are advantages in a unified and integrated public transportation infrastructure controlled and managed by one organisation, particularly in the area of the design and implementation of accessibility. Unfortunately, in many countries public transportation systems have been privatised and deregulated, leading to the involvement of a large number of organisations in different parts of the system, with associated disadvantages. However, despite the lack of an integrated approach, the use of national accessibility guidelines (for example, in the UK, see IHT 1991) and accessibility legislation has led to some degree of accessibility. However, there are still considerable variations between countries and public transport that is fully accessible to all sections of the communities has not yet been achieved.

The section is divided into two parts. Section 10.5.1 considers good design to promote accessibility, whereas Section 10.5.2 describes a number of way-finding, information and guidance systems for visually impaired and blind travellers. Details of a Bluetooth system developed for a bus transportation system in Singapore are given in Chapter 11.

10.5.1 Accessible Public Transportation: Design Issues

Public transportation systems include purpose-built buildings and other structures, such as bus and tram stops and shelters, bus and rail stations, ferry terminals and airports. There are both common and specific accessibility features for the different structures and different transportation modes.

Floors, steps, stairs and walkways

All walking surfaces, both inside and outside, should be designed to reduce glare and be non-slip even when wet. Any slopes should be shallow and pedestrian route-ways should be uncluttered and free of obstructions and street furniture.

Tactile flooring, pavement blocks and strong colour contrasts should be used to draw attention to facilities and potential hazards. The entries to steps, stairs and escalators should always be marked on all levels by tactile flooring and colour changes. Stair and escalator treads should have a colour contrasting edging strip. Open stairs and lifts should generally be avoided since they are inaccessible to people with vertigo. Any existing stairs should have handrails and banisters. Ramps and lifts should be located close to all stairs to provide an alternative route for wheelchair users. The spaces below stair flights and ramp ways should always be blocked off or handrails, raised kerbs and tactile warning floors provided if this is not possible.

Lighting

Good lighting is important for both accessibility and personal security, particularly at night, in all types of transportation centres and nodes. For economic reasons, low-pressure sodium lighting is often used but this monochromatic light has the disadvantage of colour "wash-out". A whiter light source such as high-pressure sodium lighting gives better colour definition and increases accessibility for visually impaired people. However, care has to be taken in the design of powerful lighting systems to ensure that the illumination is sufficiently diffuse and the light source adequately shaded and at a sufficient distance from users to avoid making the area inaccessible to light-sensitive people.

Information and signage

Well-designed signage and accessible information are essential to facilitate stress-free travel. The guidelines for the provision of accessible signage and information include the following (IHT 1991), slightly modified by the authors:

- Passenger entrances to stations should have a tactile map of the layout with clear and well-illuminated details of operators' names, route numbers, and destinations. Boarding point letters and numbers should be indicated by tactile, audible (preferably over a single headphone) and large print signage.
- All boarding points should be indicated by both visual and tactile signs and colour and textured floor surfaces.

- All staff should be trained to understand the mobility requirements of disabled passengers and, for instance, how to guide blind travellers to their boarding point.

Pedestrian route markers

Tactile route markers can facilitate way-finding for visually impaired and blind travellers, particularly in large transportation buildings with a complex layout. The building layout should be designed to be logical and as simple as possible to facilitate navigation and orientation. Tactile route markers use a line of tactile paving blocks to form a route through the building to important locations such as the ticket office, the toilets and the departure areas. Colour contrasting inserts with a contrasting texture can be used to indicate that a facility is close by. Japan is an example of a country where tactile route markers are widely installed. Large transportation buildings should also have tactile, audio (over headphones) and large print maps at regular locations.

Facilities

Transportation buildings generally include a wide range of facilities, such as seating areas, public waiting rooms, toilets, ticket offices, information centres, refreshment vendors and first aid points. Airports and large train stations will include other facilities, such as restaurants and bars, shops, transport police, medical facilities and chapels. Good quality visual, tactile and audio signage is important for accessibility. Plate-glass is increasingly being used, particularly in new premises. Where its use cannot be avoided, such as in viewing points, contrasting colour strips should be used to draw attention to these transparent materials. Bus shelters constructed of transparent toughened plastic also require all sides to be marked and this is frequently the case, even if only by advertising.

10.5.2 Accessible Public Transportation: Technological Information and Way-finding Systems

Making a journey by public transport involves a number of activities, including the following:

- Planning the journey and deciding how best to reach the destination.
- Using timetables or contacting service operators to obtain information about the times of services.
- Finding an embarkation point, for example, a bus stop, tram stop, train platform or aircraft departure lounge.
- Purchasing a ticket.
- Identifying and getting onto the correct bus, tram or train for the journey and finding a seat.
- Knowing when to disembark, *i.e.* the ability to identify the correct stop or station.

- Finding an exit from the station or stop and finding the correct orientation for the next stage of the journey.

Good design, including the use of tactile pavements and bold colour contrast features, can facilitate some of these activities. However, good design needs to be complemented by technological information and navigation systems. There is very little standardisation or uniformity between the different applications of active information and navigation systems in different countries and the following partial list of applications is based on reports in the literature:

- Orientation beacons to provide audio descriptions of navigation information.
- Orientation beacons to provide audio signals of important travel locations, such as ticket offices and toilets.
- Information beacons to provide audio descriptions in railway and bus stations.
- Audio information at embarkation points giving the line and destination of transport vehicles.
- Audio information inside transport vehicles with a control handset.

This list gives a flavour of some of the different types of applications, two of which are described below. However, there has been a tendency to set up isolated systems, rather than use a holistic approach to the design and implementation of information and navigation systems for blind and visually impaired travellers.

Bus identification system, USA

This system, which is distributed by the Luminator division of the US Mart IV Information Display Systems Group, uses Talking Signs® infrared transmission technology (see Section 10.4.2.1). Audio signals are transmitted using an infrared beam from a fixed beacon mounted at the front of a bus. The beam is intercepted by a hand-held receiver, which decodes the infrared signal to produce an audio message signal. The audio-signal is delivered as a voice-message to the traveller using an earphone or handset speaker. The system, illustrated in Figure 10.13, is compatible with the company's integrated voice system for embarkation point and destination announcements.

Way-finding system, Czech Republic

An extensive active information and navigation system has been implemented in Prague and several other towns in the Czech Republic. Much of the impetus for this system has come from the Czech Blind United organisation and their specialist National Centre for the Elimination of Architectural and Transportation Barriers. The achievements of this Centre include the following systems:

- Sound beacons that are remotely activated by handset controls have been located at important travel locations, including metro station entrances, escalators, railway station platforms, public buildings, banks and libraries. A photograph of a sound beacon is shown in Figure 10.14.

a b

Figure 10.13a,b. Bus identifying system (photograph reproduced by kind permission of Mark IV Luminator, USA)

Figure 10.14. Orientation audio beacon (photograph reproduced by kind permission of Czech Blind United, Czech Republic)

- Information messages that are remotely activated by visually impaired and blind travellers to identify the line number and destination of an incoming bus or tram.

10.6 Chapter Summary

The accessible environment is a recent development that is in accordance with the social model of disability. The campaigns of the disability movement, as well as other factors, have resulted in legislation, regulations and guidelines which should lead to a significant change in approach to the construction of the community environment. It is always easier to design for accessibility (as part of a design for all approach) from the start, than to modify existing structures. Therefore, it is easier to make new environments and buildings fully accessible than to modify existing buildings and environments to improve accessibility. However, the requirements of legislation, as well as other factors, are leading to changes being made in existing and older infrastructure. The chapter considered the following two main topics: the environments of streets and buildings and information and navigation technologies, as well as accessible public transportation.

The physical environment

Streets and buildings of the modern physical environment should be designed to be accessible (and easily useable) by the whole community. This has been successful to different extents. However, many of the devices and features described meet the needs of several different disability groups as well as providing improvements for non-disabled people, rather than just meeting the requirements of visually impaired and blind people. The discussion in this chapter concentrated on the special guidelines and adaptations for visually impaired and blind people. Key themes in both street and interior environments were the reduction of the hazards presented by obstacles, the use of strong colour contrasts and tactile surfaces to provide directional and information cues, and information on the presence of objects to prevent them becoming hazards.

Information, navigation and public transportation

Both route navigation and the use of public transport require information, which is frequently provided in a visual and symbolic form. The visual presentation makes this information either totally inaccessible or only partially and with difficulty accessible to visually impaired and blind people. Thus, the second part of the chapter was concerned with technological systems that make the information content of signs and notices accessible using audio signals. Systems using infrared, light, and radio-frequency technologies were described. Information, signs and notice systems that can be interrogated by the traveller may have a particular role in contributing to environmental accessibility and supporting independent travel by visually impaired and blind people.

Acknowledgement. The authors would like to acknowledge the kind assistance of staff from Talking Signs, Inc., Mark IV Luminator and Czech Blind United for the provision of and permission to publish photographs of their products and applications.

Questions

Q.1 List the main components of the accessible environment. Structure this list using a tree diagram.

Q.2 What are the different types of tactile pavement blocks called? Describe how each type is used.

Q.3 Construct a diagram of a controlled pedestrian crossing. Mark in, name and describe at least four different accessibility features relevant to visually impaired or blind pedestrians.

Q.4 List four main principles that make an exterior environment accessible to visually impaired and blind people.

Q.5 Describe the main features of audio signal environmental information systems that allow them to be located by a blind pedestrian. List the advantages and disadvantages of each of these types of system.

Q.6 List the potential applications of the Talking Signs system.

Projects

P.1 Use EU websites to determine the Framework Directives that form the basis for European actions on accessibility in the physical and urban environments. Discuss the ways in which one country in the EU has implemented these directives.

P.2 Construct an audit pro-forma that can be used to assess the accessibility of an interior environment for visually impaired and blind people. Test your form by conducting an audit of your local library for accessibility by (a) a visually impaired person and (b) a blind person.

P.3 As a continuation of Project P.2, use the Internet to obtain professionally designed accessibility audit procedures and compare the *pro-forma* constructed in Project P.2 with them.

P.4 Carry out a literature survey of recent international research, development and implementation of technologically advanced pedestrian crossings that are fully accessible to visually impaired and blind pedestrians. Evaluate your findings. This should include a discussion of the different technologies available, their advantages and disadvantages and the factors which determine the appropriate technological solution in a particular context. Make suggestions for future research directions and make recommendations for the implementation of some of the advanced technological solutions.

P.5 Develop a survey form for an accessibility audit and use it to carry an audit of your local railway station for accessibility by

(a) a visually impaired person

(b) a blind person.

References

ADAAG, 2002, Appendix A to Part 119 Americans with Disabilities Act (ADA) Accessibility Guidelines for Buildings and Facilities (ADAAG), http://www.access-board.gov/adaag/html/adaag.htm
AZBUILD, undated,
http://www.azobuild.com/details.asp?ArticleID=1310#_Weighted_Sound_Reduction

Barker, P., Barrick, J., and Wilson, R., 1995, Building Sight, Her Majesty's Stationery Office, London, ISBN 011 701 993 3

CIBSE, 2002, Code for Lighting, Chartered Institute of Building Services Engineers, London, U.K., Butterworth-Heinemann Publishers, Oxford, UK, (ISBN 0-7506-5637-9)
CIE, 1987, International Lighting Vocabulary, 4th edn. Publication CIE 17.4-1987, ISBN 3 900 734 0 70
CIE, 1995, Method of Measuring and Specifying Colour Rendering Properties of Light Sources, CIE 13.3-1995, ISBN 3 900 734 57 7
IHT, 1991, Revised Guidelines for: Reducing Mobility Handicaps: Towards a Barrier-Free Environment, The Institution of Highways and Transportation, London, U.K., ISBN 0-902-933-0-86
NI, undated, http://zone.ni.com/devzone/cda/tut/p/id/2316
Welch, P. (ed.) (1995). Strategies for Teaching Universal Design, Adaptive Environments, MIG Communications.

Resources

For more information on the work of the U.S. Access Board and to view the ADA-ABA Accessibility Guidelines, the website is www.access-board.gov

For more information on the Talking Signs® technology and its application, see the website www.talkingsigns.com

For more information and photographs on the Czech information and navigation systems, see the website www.SONS.cz

For more research orientated publications on travel by visually impaired and blind people, see the website of the Smith-Kettlewell Eye Research Institute, California Pacific Medical Centre, San Francisco, California, USA

11 Accessible Bus System: A Bluetooth Application

Learning Objectives

The flexibility and freedom offered by new wireless technologies are often discussed in the media and on the Internet. However, having this potential translated into working assistive technology systems is not so common. In this chapter, an application of Bluetooth technology for a bus alerting and information system suitable for use by visually impaired people is described. The chapter reports case study material across all the activities found in a typical prototype development project.

The chapter opens with a detailed consideration of the elements of Bluetooth technology and has a short comparison section with other competing wireless technologies. Having selected Bluetooth as the enabling technology, the chapter then reports on the design requirements for the bus alerting system. The system development is detailed along with careful consideration of the user interface needed for visually impaired people using mobile telephone technology. A discussion of future plans and commercialisation issues closes the chapter.

On completion of this chapter, the reader will be able to:

- Understand the basic working principles behind Bluetooth.
- Identify challenges faced when developing a Bluetooth application.
- Develop a simple PC-based mobile Bluetooth application.
- Appreciate the possibilities of wireless technologies in AT applications.

11.1 Introduction

Wireless technologies have caused a profound impact in many aspects of our daily lives. It has revolutionized the way we work, the way we play and the way we communicate with one another. Now, with new emerging wireless technologies, concurrent voice and data communication *via* wireless devices is now also possible. This allows short-range wireless personal communication technologies to play a key role in enabling new applications, including those in the domains of assistive technology. A useful survey of the requirements for the interconnection of assistive technology devices and information and communication technology systems has been given by Gill (2003).

This chapter will focus on an application, using wireless voice and data communication, to assist the visually impaired, physically disabled or elderly public bus commuters overcome some of the accessibility barriers when using public transportation services independently. This class of commuters faces difficulties when trying to identify the correct oncoming bus so that they can prepare themselves for boarding the bus without panic or undue stress. Thus, they often have to rely on other commuters to help in this aspect of their lives and this greatly lowers their motivation to use the public transport service freely.

A *Bus Notification System*, using Bluetooth wireless technology, can be developed to enable automatic audio notification of the requested bus number that the commuter is planning to board *via* their mobile phone or other mobile device. A group from the University of Colorado (PA 2003) has also proposed a similar application.

This chapter will thus focus on the development of a *Bluetooth Notification System for the bus commuter* as a specific example of the use of wireless technologies in assistive technology applications. The implementation, the development challenges and other possible applications of Bluetooth are discussed in the sequel.

11.2 Bluetooth Fundamentals

A brief description of the Bluetooth specifications relevant to the development of the *Bluetooth Notification System for the Bus Commuter* will be given in this section. This short review will help the reader to appreciate the constraints faced when using Bluetooth in a practical application. For a more detailed coverage of *Bluetooth Specifications*, the interested reader may refer to BSP (2005).

11.2.1 Brief History of Bluetooth

In 1994, Ericsson embarked on a project to study the feasibility of a low-power and low-cost radio interface to eliminate the use of cables linking mobile phones to their accessories. Together with four other leading companies (IBM, Nokia, Intel and Toshiba), they formed the *Bluetooth Special Interest Group*. The group, in turn, created the Bluetooth open specifications for hardware and software to promote interoperable and cross-platform implementations in 1999. The Bluetooth project soon developed into a major joint telecom and computing industrial initiative with over 2500 adopters. These include major industrial players such as 3COM, Motorola, Lucent and Microsoft.

11.2.2 Bluetooth Power Class

The current version that is most popular in applications is Bluetooth 1.1. Bluetooth operates in the globally available, licence-free ISM band of 2.4 GHz and 2.483 GHz,

Table 11.1. Bluetooth device power classes

Radio class	Power rating (mW)	Range (m)
1	100	100
2	2.5	20
3	1	10

which is reserved for general use for *Industrial, Scientific and Medical* (ISM) instruments. Depending on the power level, the effective range of Bluetooth is between 10 m and 100 m (see Table 11.1). Most of the commercial Bluetooth products work in the region of 10 m. The power source usually arises from small chemical batteries, so as not to limit the deployment possibilities of a wireless application. These low power operational modes of Bluetooth enable the battery life to be prolonged as shown in Table 11.1.

11.2.3 Protocol Stack

The Bluetooth protocol stack comprises a series of layers that allows modular development. It is similar to the *open system interconnect* (OSI) model. The stack is important to allow access to, and control of, a Bluetooth device programmatically. Stack can be viewed as driver for Bluetooth devices, similar to a graphic card or sound card driver for the Personal Computer. The layers of the stack can be realised in one of the following forms: hardware, software, firmware or a combination of these. The stack is used to implement the Bluetooth specifications. The *service discovery protocol* (SDP) advertises the stacks available in a Bluetooth device.

With the stack in place, program applications can then pinpoint the types of Bluetooth devices near one another, discover the services they offer and use the services appropriately. The layers that are relevant to this assistive technology project and that are discussed in more detail later, are as follows:

- *Host Controller Interface*. Handles communications between a host (PC) and a Bluetooth module.
- *Logical Link Control and Adaptation (L2CAP)*. Handles data from higher layers.
- *Service discovery protocol (SDP)*. Discovers services in other Bluetooth devices.
- *RFCOMM*. Provides a RS232-like serial interface.
- *Object exchange (OBEX)*. Sends and receives objects like files.

11.2.4 Bluetooth Profile

A Bluetooth profile defines a set of required functionalities for Bluetooth devices. This definition is to ensure interoperability by providing a structured set of procedure call instructions for the higher layers of the Bluetooth stack to communicate with the lower layers of the stack. As a result, Bluetooth profiles provide a way for

Bluetooth technology to be incorporated into different devices and applications and yet, still work in a unified standard way. The following profiles are used for this assistive technology project:

- *Generic Access Profile*. This is the most basic and essential profile, upon which all other profiles are built. It is used for a basic connection establishment.
- *Service Discovery Application Profile*. This profile provides the means for client applications to discover the existence of services provided by server applications as well as the attributes of those services.
- *Serial Port Profile*. This profile interacts with the RFCOMM layer in the Bluetooth stack and creates a virtual serial port (RS-232 serial cable emulation) for data transmission.

11.2.5 Piconet

Bluetooth employs a Master–Slave architecture, in which a device that initiates the connection is the master and the device that receives the connection is the slave. Piconet is a star topology net that consists of one Master and one to seven slaves. There can only be one Master in a Piconet. The role of the Master is to coordinate the communication among the Bluetooth devices in the piconet. The Master communicates with the Slave through polling; the Slaves can only transmit when they receive a "token" from the Master. The data transmission is restricted between the Master and Slave. The slave devices in the Piconet cannot transmit data directly to other slave devices in the same Piconet. Communication between the Slave devices has to be initiated *via* the Master. Piconets can be linked to form a Scatternet, a larger scale network. A Slave can participate in more than one Piconet. The intermediate Slave between two Piconets will serve as the bridge node for communication.

11.2.6 Bluetooth and Competing Wireless Technologies

In this section, Bluetooth will be compared briefly with other common and existing wireless communication standards. Table 11.2 shows the comparisons made between the main features of the IrDA Infrared and Wireless LAN.

Table 11.2. Comparisons between key features of infrared, 802.11b(g) and Bluetooth 1.1

Parameter	Infrared	802.11b(g)	Bluetooth 1.1
Medium	Optical/unidirectional	RF/omnidirectional	RF/omnidirectional
Gross data rate (Mb/s)	1.152/4/16	5.5/11/55[a]	1[b]
Max range	20 cm/1.2 m	100 m	10 m/50 m/100 m
Data and voice (real-time) communication	Data only	Both	Both

[a] Data rate of up to 55 Mb/s is expected for 802.11g
[b] Greater data rates of up to 2 Mb/s are expected for future specifications of Bluetooth

11.2.6.1 IrDA Infrared

Infrared data communication is unidirectional and has a short nominal range of about 10 m only. Communication between mobile devices such as personal data assistants (PDAs) using infrared technology is limited to areas within a line-of-sight as IrDA infrared cannot penetrate walls and furniture. The advantage of Bluetooth over Infrared is that it does not have this line-of-sight constraint, since a Bluetooth transmission is omnidirectional. Due to Bluetooth's greater nominal range of up to 100 m, Bluetooth transceivers do not have to be located near to each other for data transmission as compared to infrared transceivers.

11.2.6.2 IEEE 802.11b and 802.11g

Commonly known as Wireless LAN, 802.11b serves data transmission of up to 11 Mb/s, while 802.11g is extended from 802.11b and serves data transmission of up to 54 Mb/s. Despite the high data transmission of Wireless LAN, it requires too much power from the mobile phones and the high transfer rate is unnecessary for small data transfer applications of mobile phones. Thus, in this project, Bluetooth is a more suitable wireless technology as it consumes far lesser battery power. Furthermore, the amount of data transfer in the project is not sufficiently large to warrant the use of Wireless LAN.

11.2.6.3 New Emerging Wireless Technologies

Zigbee, UWB (ultra wide band) and NFC (near field communication) are some of the new emerging short-range data transfer wireless technologies. From the application point of view, they differ mainly in term of data throughput, transmission distance and power consumption. In term of support and availability, those given for Bluetooth are much more extensive than these new emerging protocols. Some of these new wireless technologies may also be used to complement Bluetooth, for example, the connection initialisation phase of Bluetooth (approximately 3 s) can be reduced to milliseconds using NFC (NFC 2005).

11.3 Design Issues

The target application of *Bluetooth Notification System for the Bus Commuter* is to overcome any barriers bus commuters may have in identifying an approaching bus to determine whether it is a service that goes to their desired destination and is therefore suitable for boarding.

Those who are visually impaired, physically disabled or elderly will often experience this barrier to independent use of the public transport system. This identification or notification system can also be extended to those who are facing difficulty in using other public transport systems, such as taxis. Under the current transport system, such a commuter may have difficulty in flagging down an approaching bus on time. At the same time, a visually impaired commuter or an elderly person with poor vision is often unable to determine if the right bus is approaching them. In addition, if a visually disabled commuter has boarded the bus, they may have difficulty in determining when the bus has reached their desired destination. These commuters often have to rely on fellow commuters around them for guidance. Thus, their ability to travel independently is greatly hampered.

11.3.1 System Architecture

A solution based on Bluetooth technology is proposed to remove these particular accessibility barriers.

The proposed system as shown in Figure 11.1 consists of three Bluetooth components:

1. A Bluetooth embedded system installed on the bus
2. A Bluetooth server module located at the bus stop
3. A Bluetooth-enabled phone carried by the commuter pre-installed with Java application software

When the user or commuter approaches the bus stop, a connection will be established between their mobile and the bus stop server *via* Bluetooth connection. Each bus stop is associated with a unique identification (ID) code. The bus stop server will notify the user with the ID of the bus stop. Other information (bus route, frequency *etc.*) can be accessed from this server as well.

Via the mobile phone, the user will key in the bus number that they are waiting for, as shown in Figure 11.2. Upon receiving this message, the bus stop server will, *via* Bluetooth, detect the bus once it is within the Bluetooth messaging range and transmit a notification message to the user's mobile phone to alert the user that the bus they require is coming. An audio message is then played to inform the user that

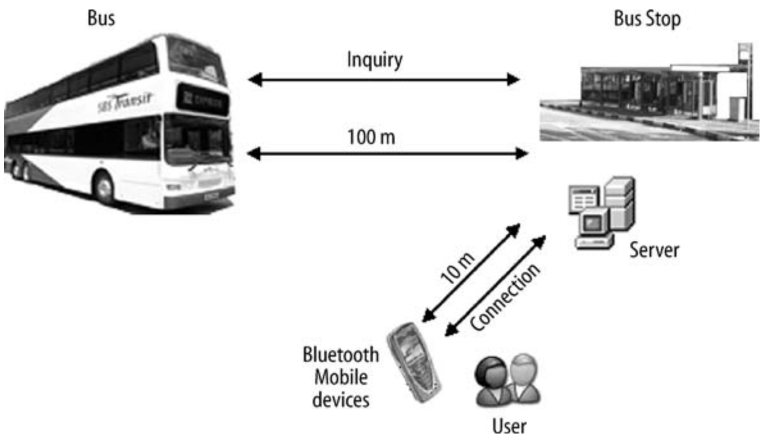


Figure 11.1. Proposed system architecture

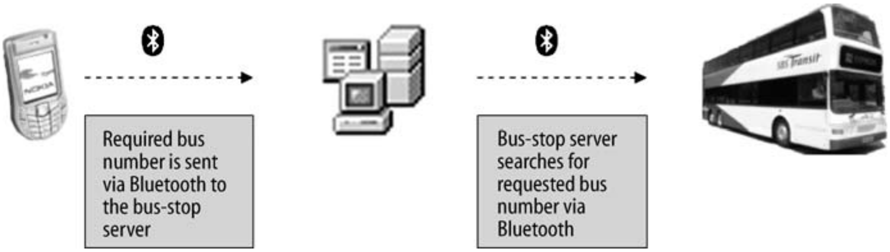


Figure 11.2. Bus commuter keys in bus number that they are waiting for into their mobile phone

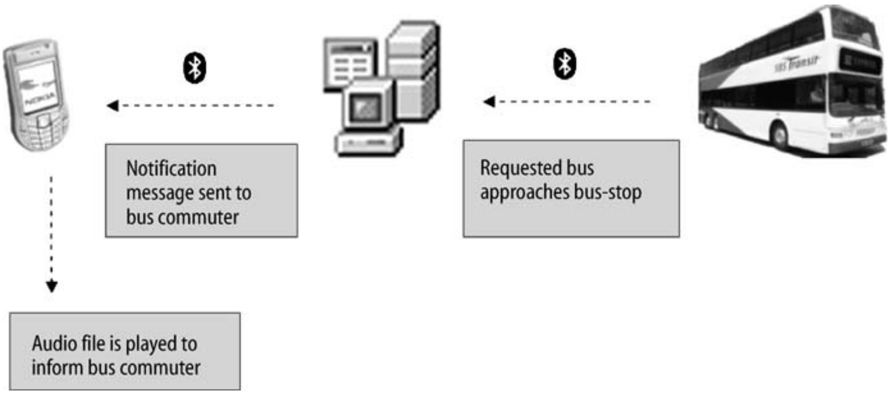


Figure 11.3. Bus commuter receives notification message that the requested bus is approaching

their bus is now approaching, as illustrated in Figure 11.3. Through the Bluetooth system on the bus, the bus driver can also be notified of commuters waiting for it at the bus stop so that the driver will stop and collect them.

On the other hand, when the commuter (who has boarded the bus already) wants to alight from the bus at a specific bus stop, they can enter their desired destination bus stop ID into their mobile phone. Once the bus has reached the designated bus stop, the bus driver is alerted and is able to let the commuter alight.

11.3.2 Hardware Requirements

Based on the proposed architecture, a hardware prototype was developed. The list of hardware used is given in Table 11.3.

11.3.3 Software Requirements

A suitable software development kit is required to program the Bluetooth protocol stack *via* the USB Bluetooth adaptor, which is to be connected with the PC server.

Java Bluetooth API, or JSR-82 (JAVA 2005), is used as the preferred implementation of the Bluetooth protocol stack. The key reasons for this preferred choice are

Table 11.3. List of hardware used in the project

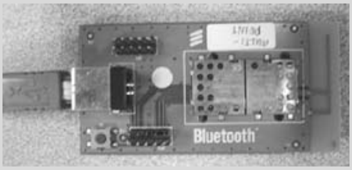
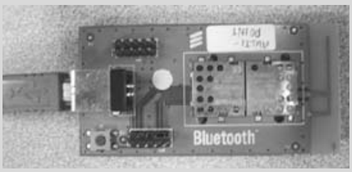
Components	Description
PC Server (*Bluetooth Server*) 	The PC server that waits for client (mobile phone) connection to send the requested bus number
Bluetooth Adaptor (*Bluetooth Server*) 	For Bluetooth connectivity, the PC server needs to be connected to the USB Bluetooth adaptor. The adaptor provides the radio hardware for proper Bluetooth connection
Ericsson Bluetooth Radio Kit (*Bluetooth Server*) 	The Ericsson Bluetooth radio kit is used to poll for Bluetooth addresses of devices in the vicinity
Mobile Phone (*Bus commuter*) 	A phone that supports JSR-82 (Java Bluetooth API) and JSR-135 (Java Multimedia API) for Bluetooth connection and to play audio notification when the requested bus approaches the bus-stop. The models used in the project were Nokia 6630 and 6230
Ericsson Bluetooth Radio Kit (*Bus*) 	The Ericsson Bluetooth radio kit gives the unique Bluetooth address present in each bus. This allows the Bluetooth bus-stop server to identify the requested bus number

Table 11.4. List of Bluetooth software development kits. Blue Cove (*highlighted*) was chosen for the project

Company	Java platforms	Operating systems	Price
Atinav	J2ME, J2SE	Win-32, Linux, Pocket PC	US $ 4999 for 10 licences
Avetana	J2SE	Win-32, Mac OS X, Linux, Pocket PC	€ 25 for 3 licences for Win-32, Free for Linux
Blue Cove	**J2SE**	**WinXP SP2**	**Free**
JavaBluetooth.org	Any platform that supports javax.comm	Win-32, Mac OS X, Linux	Free
Rococo	J2ME, J2SE	Linux, Palm OS	€ 2500, free with limited functions for students
Bluez	J2SE, C	Linux	Free

that JSR-82 is independent of the manufacturer's Bluetooth radio design and it is independent of operating platform.

This implies that it has a standard, fixed set of APIs that is non-proprietary as long as the Java Virtual Machine is installed. However, it is known that, this has not turned out to be completely true for applications that run on different makes of devices. Different device vendors will have slightly different implementation of mobile information device profile (MIDP) (Yuan 2003).

Currently, there are a number of established Java Bluetooth protocol stacks for various operating systems. Table 11.4 shows some of the more common stacks and the corresponding programming languages to access the stacks.

BlueCove, an open-source Java Bluetooth stack for Microsoft Windows XP Service Pack 2 as illustrated in Figure 11.4, was chosen due to its relatively free licence (under open source agreement). Microsoft Bluetooth API stack is incorporated in Windows XP SP2. This stack is a socket-style (WinSock) C API that allows any C-based applications to operate a generic USB Bluetooth adaptor. However, there

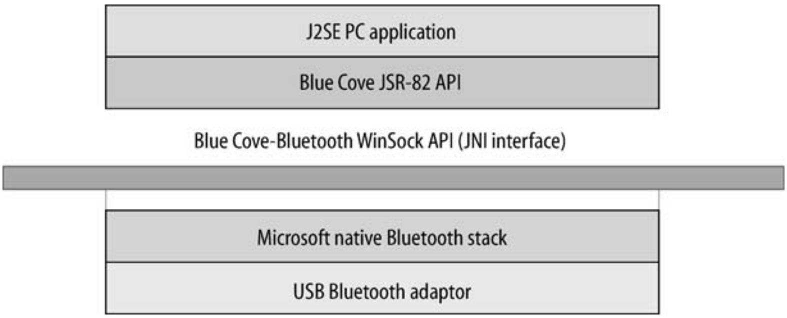


Figure 11.4. Architecture of Blue Cove in a Microsoft Windows environment

is currently no Java version of this API. Hence, BlueCove was developed to bridge this gap by developing a JSR-82 API on top of Microsoft Bluetooth API. This integration was done through Java Native Interface (JNI). This technology allows Java applications to use the Microsoft Bluetooth native C-based library.

As BlueCove is based on JSR-82 API, it offers a very open and friendly environment for programming. Being an open-source, non-commercial software, Blue Cove has its limitations in that it does not support many Bluetooth profiles. Currently, it supports only the *Generic Access Profile* (partially), *Service Discovery Profile* and *Serial Port Profile*. However, in this project, BlueCove is sufficient for the development of the prototype.

In addition, the Ericsson Bluetooth radio kit is used to develop the server application to poll for Bluetooth addresses in the vicinity (see Figure 11.11).

This is used in coordination with the JSR-82 application developed with BlueCove to search for the requested bus number as defined by the bus commuter.

11.4 Developmental Issues

The realisation of the Bluetooth notification system prototype is illustrated in Figure 11.5. There are three components:

- Bus stop Bluetooth server attached to a USB, running the BlueCove Java Bluetooth stack.
- Bluetooth mobile phone client capable of running Java (J2ME), Java Bluetooth API (JSR-82) and Multimedia API (JSR-135).
- Bluetooth-bus as simulated by the Ericsson Bluetooth Radio Kit.

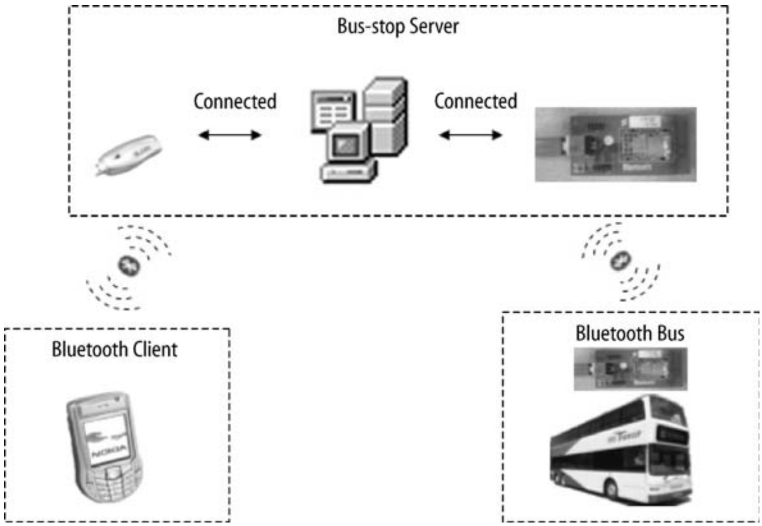


Figure 11.5. Implementation of system prototype

As shown in Figure 11.5, the Bluetooth server is started and ready to accept a client connection. The bus commuter uses their mobile phone to initiate a Bluetooth connection with the bus stop server as seen by the personal computer with the USB Bluetooth adaptor. Once a Bluetooth connection has been established, the bus commuter will key in the required bus number into their mobile phone using the preinstalled Java (J2ME) software. The number is then sent to the server.

Using the requested bus number, the server will initiate the Ericsson Bluetooth Radio kit to search for the requested bus number, based on its corresponding Bluetooth address.

Once the requested bus is within the range of the Ericsson kit, a match between the Bluetooth address and the bus number will be made, and if there is a match, a notification message will be sent to the Bluetooth client. This, in turn, activates the audio notification message of the Bluetooth client, which informs the bus commuter that their bus is approaching.

11.4.1 Bluetooth Server

In the project, as the *Serial Port Profile* is used, the server needs to define a service record to track the *Serial Port Profile* service. Thus, the particular service of the server is advertised to any client who wants to use the particular service.

After registering the service record in the service discovery database (SDDB), the server application initiates a blocking call on its incoming Bluetooth connection port. It waits for a client application to initiate contact with the server to access the service. In this case, the mobile phone (client) of the bus commuter will try to initiate contact with the server to access the advertised service to send the requested bus number to the server. The client application and the server application then establishes a Bluetooth connection.

Once the data transmission completes (server receive requested bus number), the client disconnects from the server. The server then goes back to the state where it waits for the next client connection. The flowchart, depicted in Figure 11.6, illustrates the server application.

Note that the client will disconnect from the server once the requested bus number has been transmitted. This will free the server for other user connections. A more elegant method would be to perform a master/slave switch and maintain the connection. This situation is similar to the LAN access point application; whereby the connecting devices join a piconet by forming a small piconet containing themselves as master and server as slave, after which a master/slave switch to allow the server to control the piconet. This method saves resources as it is a waste of resources for the server to constantly poll devices for connection (Bray and Sturman 2002). However, not all the mobile devices support the master/slave mechanism. Thus, a disconnection is implemented to allow other user connections.

11.4.1.1 Processing Client Connection

Once the Bluetooth connection is established between the server and the client, the requested bus number is sent from the client to the server *via* a stream connection

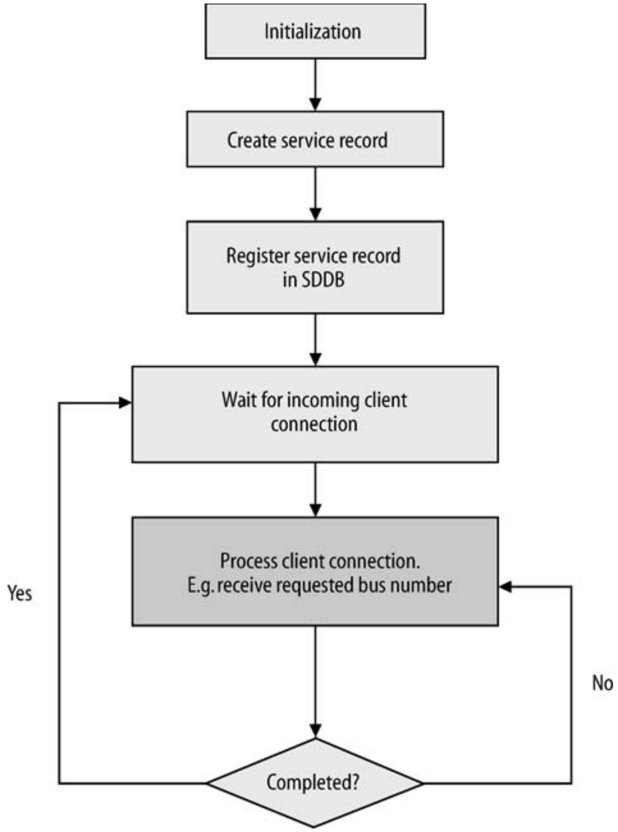


Figure 11.6. Flowchart of the initialization of server

using the *RFComm* layer of the *Serial Port Profile*. The requested bus number, in this case “96”, is sent to the server. An acknowledgement message from the server is sent to the client. The client replies with a “Waiting” message and waits for the notification message that the requested bus number is approaching the bus stop.

In the meantime, the Ericsson Bluetooth radio kit application starts to poll for any available Bluetooth addresses in the vicinity of the server. The server contains a list of numbers of buses that will stop at the bus stop and their corresponding Bluetooth addresses. For every Bluetooth address detected by the Ericsson radio software, the server will check if it corresponds to the requested bus “96” in the server database. Once a match has been found, a notification message is sent to the client and the client plays an audio notification to alert the bus commuter that their bus is approaching. The flowchart, shown in Figure 11.7, illustrates the concept.

The screenshot in Figure 11.8 illustrates the server finding a matching Bluetooth address, 000AD9B862B0, which corresponds to the bus number “96”. This implies that the requested bus number is approaching the bus stop. A notification message

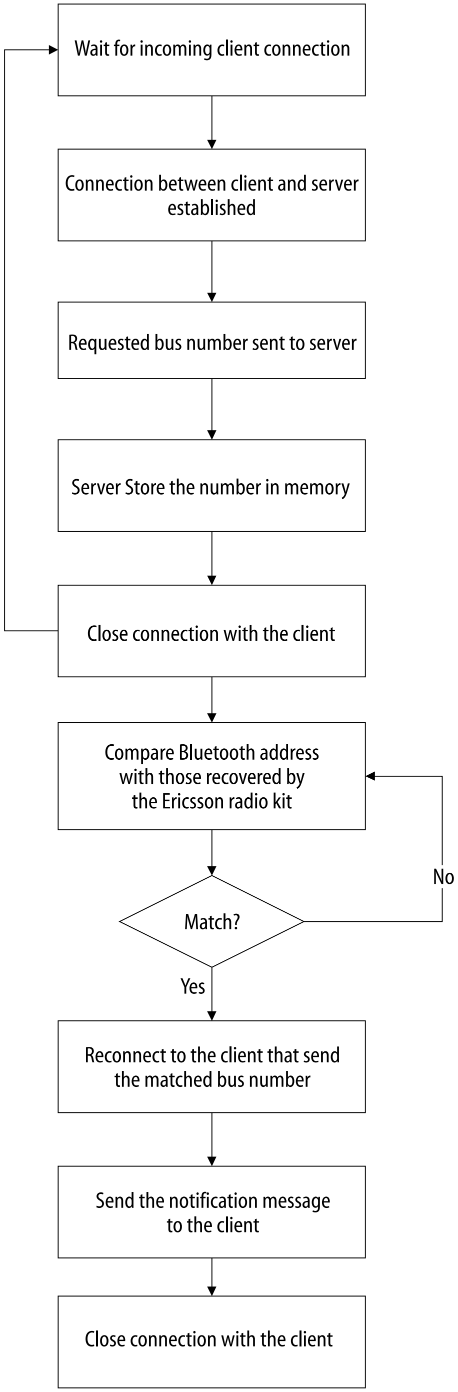


Figure 11.7. Flowchart of server polling to find a matching Bluetooth address

Figure 11.8. Screenshot of server application when the requested bus is found

is then sent to the client as shown upper dashed box of Figure 11.8. Finally, the server closes the Bluetooth connection with the client and waits for the next client to connect (lower dashed box).

11.4.1.2 Ericsson Bluetooth Module

Two Ericsson Bluetooth modules are used in this project to detect the number of the approaching bus. A single-point and multi-point Bluetooth module is to be mounted on the bus and the server respectively. The module on the bus will periodically enter the inquiry scan state to listen for an ID packet transmitted from the inquirer (Bus stop server). During the inquiry procedure, the bus mounted module will respond with a frequency hop synchronisation (FHS) packet to the inquirer. Based on the FHS packet received, the server will be able to determine the Bluetooth address (BD_ADDR) of the approaching bus module. The unique BD_ADDR is then translated to the bus number. In this method, no connection needs to be established to transmit the bus number information, thus reducing the overall response time of the system.

The proprietary Ericsson *PC Reference Stack* allows the control of the Bluetooth module at a different stack layer. In the initialization process, the application configures the HCI layer and registers with the stack connection manager (SCM) component.

Figure 11.9 show the address of the bus's Bluetooth module detected by the server.

The HCI_ReqPeriodicInquiry function is used to set the module into periodic inquiry stage. The HCI_ReqPeriodicInquiry will return the address of the bus Bluetooth module detected in the vicinity.

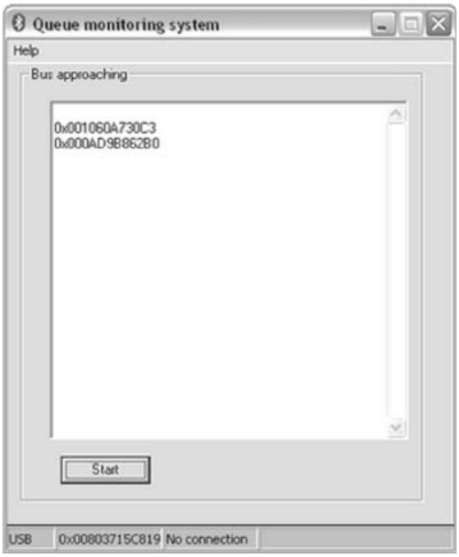


Figure 11.9. The address of the bus's module detected by the server

11.4.2 Bluetooth Client (Mobile Device)

For the client application, there are five main components:

- Initialization of the client device.
- Device discovery.
- Service discovery.
- Connection and notification.
- User interface.

11.4.2.1 Initialization of Client

The client (mobile devices) needs to be initialized first. This means that the service address (Serial Port Profile service), provided by the bus stop server, is incorporated in the client application. This is to enable the client to search for the correct Bluetooth server and the correct service to initiate connection.

11.4.2.2 Device Discovery

Once the client device is initialised, it starts to carry out device inquiry for the available Bluetooth devices in the vicinity. For every Bluetooth device found, the corresponding name and Bluetooth address of each device is stored in a hash table in the client application. Once the process of device discovery has completed, the client application moves to the next stage—service discovery. The flowchart of this

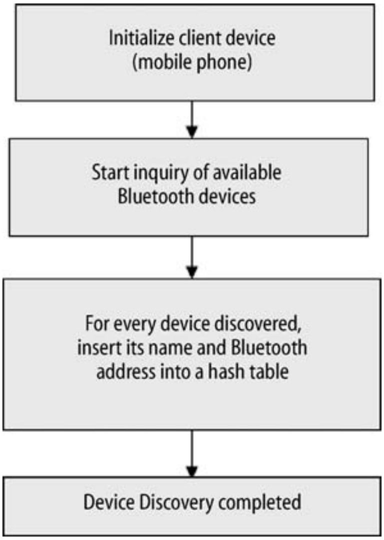


Figure 11.10. Flowchart of device discovery process of client device

Figure 11.11. Screenshot of bus commuter mobile phone that displays the results of the Device Discovery process

process and the screenshot of the client device are shown in Figures 11.10 and 11.11, respectively.

From the screenshot in Figure 11.11 it can be observe that the client device has detected a total of two Bluetooth devices in the vicinity.

11.4.2.3 Service Discovery

After the device discovery process has completed, the service discovery process begins. All the available Bluetooth devices data are now stored in the hash table. The client application will search for each device data: the Bluetooth address, the type of services and corresponding service address offered by each device.

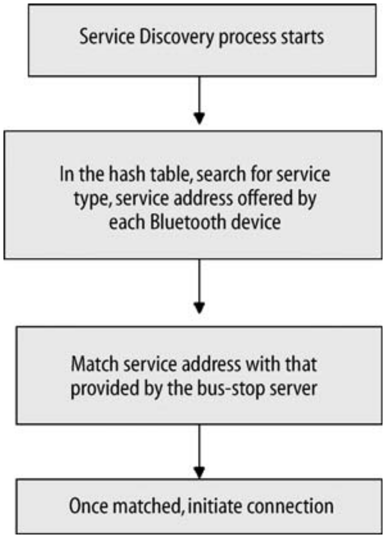


Figure 11.12. Flowchart of service discovery process of client device

If the Bluetooth and service addresses correspond to that of bus stop server, the client device will attempt to initiate a Bluetooth connection with the remote device (which is the bus stop server in this case) as shown in the flowchart of Figure 11.12.

11.4.2.4 Connection and Notification

Once a match between the service address of the remote device and that given by the bus stop server is found, the client device will attempt to initiate a connection to the bus stop server. As soon as the connection is successful, the requested bus number keyed in by the bus commuter will be sent *via* Bluetooth stream connections to the server (Figure 11.13). The server will acknowledge by sending a "RECEIVED" message back to the client to inform it that it has received the requested bus number. The client then waits for the server to check if the requested bus is approaching the bus stop (Figure 11.14).

Once the requested bus is approaching the bus stop, the server will send an "ARRIVED" notification message to the client *via* Bluetooth to indicate that the requested bus is arriving. The client will read in the message and will play an audio file to alert the bus commuter that their bus is approaching as depicted in Figure 11.15.

11.4.3 User Interface

An assistive feature of the Bluetooth bus notification system is that it incorporates an audio interface for the client application (in the mobile phone). This is especially

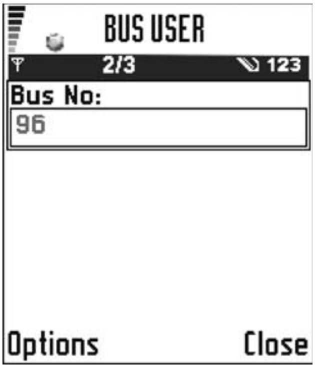


Figure 11.13. Screenshot showing that requested bus number is 96

Figure 11.14. Screenshot showing that the bus stop server has received the requested bus number and the client is waiting for the bus

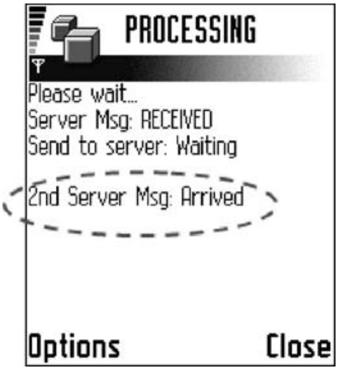


Figure 11.15. Screenshot showing that requested bus number is approaching the bus-stop

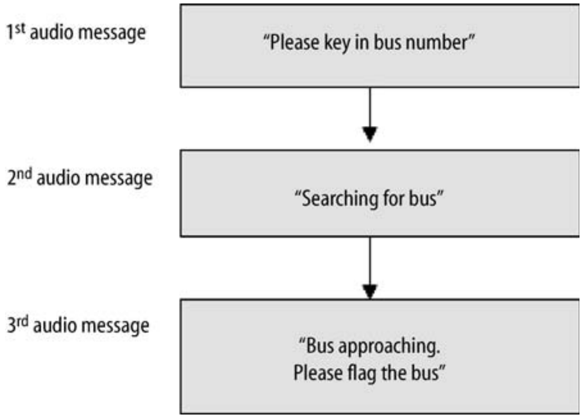


Figure 11.16. Flowchart demonstrating the sequence of audio files for the client application

important for the visually impaired commuter. The audio interface comprises a number of audio (mp3) files are stored in the client application. They are played at various stages of the client application as the bus notification procedure occurs.

As seen in Figure 11.16, the first audio file will sound to prompt the bus commuter to enter the bus number that they wish to board and want to be notified as it approaches the bus stop. Once completed, the second audio file will instruct the commuter to wait for the client application to connect. The final audio file will play once the requested bus service is approaching the bus stop.

11.5 Commercialisation Issues

According to statistics given by the Singapore Association of the Visually Handicapped (SING 2003), there are 2439 visually impaired people registered with this organisation as of March 2003. The proposed application, *Bluetooth Notification System for the Bus Commuter* may not be economically viable if it only caters for a very small community group. To make the proposed system more economically feasible, the solution is to enhance the overall system framework to be beneficial to the wider group of bus commuters and incorporate other features of value to the operation of the bus companies.

The proposed system requires all the buses and bus stops to be equipped with a Bluetooth module and this indirectly increases the normal operating cost of the public bus transport. Therefore, there is a need to expand the use of the Bluetooth module for additional features to justify the investment costs. For example, a possible use of the Bluetooth modules is to provide wireless diagnostics and conditions-based maintenance support for the bus fleet (MIL 2003). Thus, when the bus returns to the terminal after daily service, the maintenance data are wirelessly transmitted *via* Bluetooth to the server in the depot and bus maintenance schedules can be optimized.

The new generation of mobile phones are now all equipped with Bluetooth technology, so that no additional cost is incurred on the user's side in the near future. Thus, the mobile phone functional requirements used in this project will all be available and a bus notification system for wide community access to the public transport system is perfectly feasible at no significant incremental cost to the bus commuter.

11.6 Chapter Summary

In this chapter, the use of wireless technologies to construct an assistive technology application was discussed. Among the short-range wireless technologies, Bluetooth has the highest market penetration and support. Through the specific implementation of the *Bluetooth Notification System for the Bus Commuter* project, the principles and mechanisms of Bluetooth were explored in the chapter.

The aim of this assistive technology project was to develop a Bluetooth-based bus notification system that could remove information accessibility barriers to the independent use of the public bus system experienced by community groups who are visually impaired; such groups include the blind, visually impaired and the elderly commuters.

The chapter opened with a review of Bluetooth fundamentals, including a brief history of Bluetooth before discussing the structure of Bluetooth and its related tools. This review also briefly examined some competing wireless technologies to establish those singular advantages of the Bluetooth paradigm.

A major portion of the chapter was devoted to the engineering of this Bluetooth assistive technology application. The complete sequence of design and constructional steps was described. These included the system architecture, the hardware and software requirements followed by the construction of the Bluetooth development, namely the Bluetooth server, the Bluetooth client (Mobile device) processing operations and the audio user interface. The chapter closed with a brief look at some of the issues relevant to moving from prototype system to widespread and general introduction in the public bus system.

The prototype bus notification system presented convincingly demonstrates how Bluetooth has the potential to be used in enhancing information accessibility in a public transport system.

Questions

Q.1 List some assistive technology applications that use short-range wireless technologies. Compare the merits of the wireless technology used with other possible wireless technologies to realise the assistive technology application.

Q.2 Describe the differences between the Bluetooth stack and Bluetooth Profile, highlighting the stack and profile needed for audio transmission *via* Bluetooth.

Q.3 What are the functions of the service discovery database (SDDB)?

Q.4 Describe the procedures for Master/Slave switch mechanism. Discuss the scenarios under which the Master/Slave Switch is employed. Compile a list of mobile devices that support the Master/Slave Switch.

Q.5 During the inquiry phase, what information is exchanged between the inquirer and the target devices?

Q.6 Suggest other applications that can make use of the Bluetooth module mounted on the bus.

Projects

P.1 In the bus notification project described in this chapter, a PC was used as the server located at the bus stop. Implementing the functionality of the server in an embedded system would improve the portability and reduce the cost of the server significantly. Hopkins and Antony (2003) have given examples of the use of the wireless Bluetooth embedded system Micro bluetarget. Investigate and consider the implementation of the server in a wireless embedded system.

P.2 Currently, some restaurants have implemented Bluetooth customer reservation systems. There is also a trend for restaurants to adopt wireless order taking. For a wide range of reasons, a system that allows the customer to download the menu and place an order *via* mobile devices will be useful for both customers and restaurant operators. The techniques described in this chapter would allow the visually impaired to utilise such a system. Consider all the enduser and engineering aspects of developing an accessible menu ordering system based on Bluetooth technology.

P.3 Bluetooth group—this project makes use of the JSR82 enabled mobile devices to create software that will form a virtual link list between all the participating mobile devices. The JSR82 API allows the software developed to control the Bluetooth modules in the mobile devices. The objective of this project is to establish an ordered link list Bluetooth connection between the mobile devices held by users moving in a group convoy movement. The purpose is to detect any loss of member from the group and alert the group leader.

The group leader will first activate the software as head. This will cause the mobile device to send out a Bluetooth inquiry signal to look for participating members. The second driver in the convoy will activate the software as member and accept connection request of the leader. After which the mobile device of the second member will send out the inquiry signal to look for the third member's mobile devices, and the inquiry goes on until the linked list is formed.

In Bluetooth protocol, what obtains is an alternate master and slave relation:

$$\underset{\text{(User 1)}}{\text{Master}} \rightarrow \underset{\text{(User 2)}}{\text{Slave|Master}} \rightarrow \underset{\text{(User 3)}}{\text{Slave|Master}} \rightarrow \underset{\text{(User 4)}}{\text{Slave|Master}} \rightarrow \underset{\text{(User5)}}{\text{Slave|Master}}$$

When a member detaches from group, the leader will be alerted about the member who is missing. This application can be useful for group movement coordination, *e.g.* group cycling, vehicle convoy and group outing for community groups who have visually impairments.

References

Bray, J., and Sturman, C.F., 2002, Bluetooth 1.1: Connect Without Cables, 2nd Edition, Prentice Hall, New Jersey

BSP, 2005, Bluetooth Specification Documents, https://www.bluetooth.org/spec/

Gill, J., 2003, Requirements for the Interconnection of Assistive Technology Devices and Information and Communication Technology Systems, http://www.tiresias.org/reports/inter.htm

Hopkins, B., and Antony, R., 2003, Bluetooth for Java, 1st Edition, Apress

JAVA, 2005, JSR 82: JavaTM APIs for Bluetooth, http://www.jcp.org/en/jsr/detail?id=82

MIL, 2003, Bluetooth in Public Transportation Bus Fleet of Milan, Italy, http://whitepapers.zdnet.com/casestudy.aspx?dtid=3&scid=1121&docid=113218

NFC, 2005, About NFC Technology, http://www.nfc-forum.org/aboutnfc/about_nfc_technology/

PA, 2003, Palo Alto, HP and Partners Build on HP's Pocket PC Platform to Bring Assistive Technology Advances to People with Disabilities, http://www.hp.com/hpinfo/newsroom/press/2003/031015a.html

SING, 2003, Singapore Association of the Visually Handicapped Annual Report, http://www.savh.org.sg/publications/annual_reports/2003/swd_report.pdf

Yuan, M.J., 2003, A comparison of the philosophies, approaches, and features of J2ME and the upcoming .Net CF, http://www.javaworld.com/javaworld/jw-02-2003/jw-0221-wireless.html

12 Accessible Information: An Overview

Learning Objectives

Increasingly complex media and technology are being used to transmit information. However, a degree of familiarity with new technological developments is required to make the best use of information media and technologies, so it is important that this area remains accessible to the visually impaired and blind community. It should be noted that there are currently many people who either do not have access to modern communications and information technology or do not wish to use it. It is therefore important that information continues to be provided in other lower tech and more traditional ways.

This chapter opens with a review of the principles and technologies of low vision aids that are used to access print. Sections on audio transcription and Braille as access routes to print information then follow. However, it is the recent developments in speech processing and speech synthesis technology that are the drivers of the wider use of audio as an information interface for visually impaired and blind people. Major sections of the chapter describe the accessible computer and the accessible Internet. Both are extremely important in the processing and provision of information and there are many interface options to make these systems accessible to visually impaired and blind people. Finally, since telecommunications is an important information channel and mobile telephony is increasingly obtaining computer and Internet capabilities, the chapter closes by reviewing accessible telecommunications technology.

The learning objectives for the chapter are:

- Understanding a taxonomy for low vision aids.
- Appreciating the assistive technologies that are used for the audio and tactile transcriptions of print materials.
- Understanding of the different input and output devices that can be used to make computer systems accessible.
- Understanding how the World Wide Web can be made accessible to visually impaired and blind people.
- Studying technologies for accessible telecommunications.

12.1 Introduction

Access to information is becoming increasingly important and the term *Information Society* is often used, with a particular stress on electronically transmitted information. Most of this information is obtained *via* the visual and auditory senses and therefore, unless it is available in alternative formats and/or appropriate technology is available to make it accessible, people with sensory impairments will be unable to access a large part of this information.

Electronic transmission by information and telecommunications technologies has become extremely important in the industrialised countries. The term *digital divide* has been used to describe the gap between those who do and do not have access to computer technology and the Internet in particular and it has been suggested that those on the wrong side of the digital divide will be the new *have-nots*.

Access to computer technology is generally through a graphical user interface and therefore not easily accessible to blind people. Screen reader technology (see Section 12.6 and Chapter 13) is fairly well developed, but can only be used if documents are appropriately designed. The issues are particularly complex in the context of the World Wide Web with its multimedia, multimodal potential for the presentation of information using text, speech, animations, photographs, video, colour effects and sound effects. However, as well as complicating accessibility, this multimedia potential also enables the Web to be made accessible. This requires attention to the accessibility of both web authoring tools and web content. As will be discussed in Section 12.7, guidelines for both areas have been drawn up by the World Wide Web Consortium (W3C) Web Accessibility Initiative.

In the area of telecommunications, simple accessibility features can be used to enable blind and visually impaired people to use telephone handsets. Voice dialling and voice caller identification are also available. The increasing popularity of '*texting*' or sending short messages in text form using mobile telephones requires more sophisticated technological solutions based on screen readers. An overview of accessible telecommunications for blind and visually impaired people is presented in Section 12.8.

Despite the increasing focus on information and communication technologies, print media continue to be an important and frequently used means of conveying information. In addition to text, print media include tables, graphs, pictures, photographs, maps, musical notation, as well as several other types of graphical representations. Access to the different types of print media for blind and visually impaired people raises both common and distinct issues. These issues and the associated technologies will be discussed in this chapter and in more detail in Chapters 13–16.

An important and frequently forgotten group of technologies are the so-called 'low vision' aids. They are designed for blind and visually impaired people with some useful vision. Low vision aids to support access to print media will be discussed in Sections 12.2 and 12.3. The chapter will then continue in Sections 12.4 and 12.5 with an overview of information access based on two of the main sensory modalities blind and visually impaired people use to access information, hearing

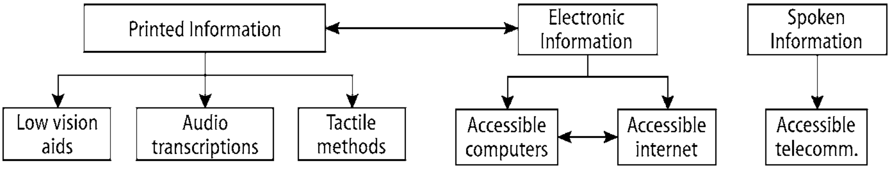


Figure 12.1. Classifying information access technologies

and touch. Sections 12.6–12.8 consider access to information and telecommunications technologies, with the focus on hardware in Section 12.6 and software and, in particular, the design of accessible web content and using web authoring tools, in Section 12.7. Accessible telecommunications for blind and visually impaired people are considered in Section 12.8. An overview of these information access technologies is presented in Figure 12.1.

12.2 Low Vision Aids

Low vision aids is the term used for a group of assistive technology devices designed for visually impaired and blind people with some useful vision. However, despite being an important group of assistive technology systems that could be of benefit to a large number of blind and visually impaired people, they are often overlooked. For instance, there are 1.7 million people in the UK who are unable to read standard print with any ease and some 36% of blind people can read large print (Gregory 1996). Thus, many people who are registered as blind have some useful visual faculties that could be augmented by an appropriate low vision aid to, for instance, give them access to print and mobility.

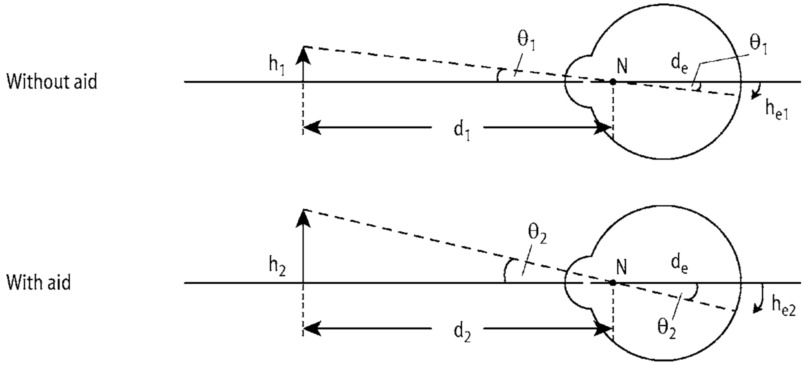
12.2.1 Basic Principles

The technology of optical low vision aids is well developed. However, there may be other technological approaches that could be developed to extend the range of applications and/or users. The aim of this section is to provide an overview of currently available low vision assistive technology systems. The interested reader is referred to Dickinson's book (Dickinson 1998) for a comprehensive and detailed presentation of low vision technology. Dickinson's taxonomy of low vision aids will be followed in this section.

Currently available low vision assistive technology is based on optical systems using magnification. Figure 12.2 illustrates the impact of a low vision aid by showing the image on the retina with and without the aid that increases the image size on the retina.

Since the retinal image size depends on the angle θ subtended at the nodal point of the eye, the magnification obtained by a low vision aid is given by

$$\text{Magnification, } M = \frac{\text{Retinal image size with aid}}{\text{Retinal image size without aid}} = \frac{\theta_2}{\theta_1} \tag{12.1}$$



Key	Without aid	With aid
Height of object	h_1	h_2
Distance of object from N	d_1	d_2
Angle subtended by object	θ_1	θ_2
Height of image at retina	h_{e1}	h_{e2}
Distance of image at retina from N	d_e	d_e
Eye node position	N	N

Figure 12.2. Low vision aids: generic analysis

Useful formulae are obtained by making the following approximations for small angles θ_1 and θ_2 (see Figure 12.2):

$$\begin{aligned} \theta_1 &\cong \tan\theta_1 = \frac{h_1}{d_1} \\ \theta_2 &\cong \tan\theta_2 = \frac{h_2}{d_2} \end{aligned} \tag{12.2}$$

where the symbols θ, h and d denote the angle subtended by the retinal image at the nodal point of the eye, the size of the object and the viewing distance respectively. Subscripts '1' and '2' denote *without aid* and *with aid* respectively. This gives the following expression for the magnification, M:

$$\text{Magnification, } M = \frac{\theta_2}{\theta_1} = \frac{\tan\theta_2}{\tan\theta_1} = \frac{\dfrac{h_2}{d_2}}{\dfrac{h_1}{d_1}} = \frac{h_2 d_1}{h_1 d_2} \tag{12.3}$$

The following four different cases, illustrated in Figure 12.3 and which lead to different types of assistive technology, are distinguished by Dickinson (1998):

1. Increasing the object size, but keeping the viewing distance constant
2. Decreasing the viewing distance while keeping the object size constant
3. Creating a magnified real image

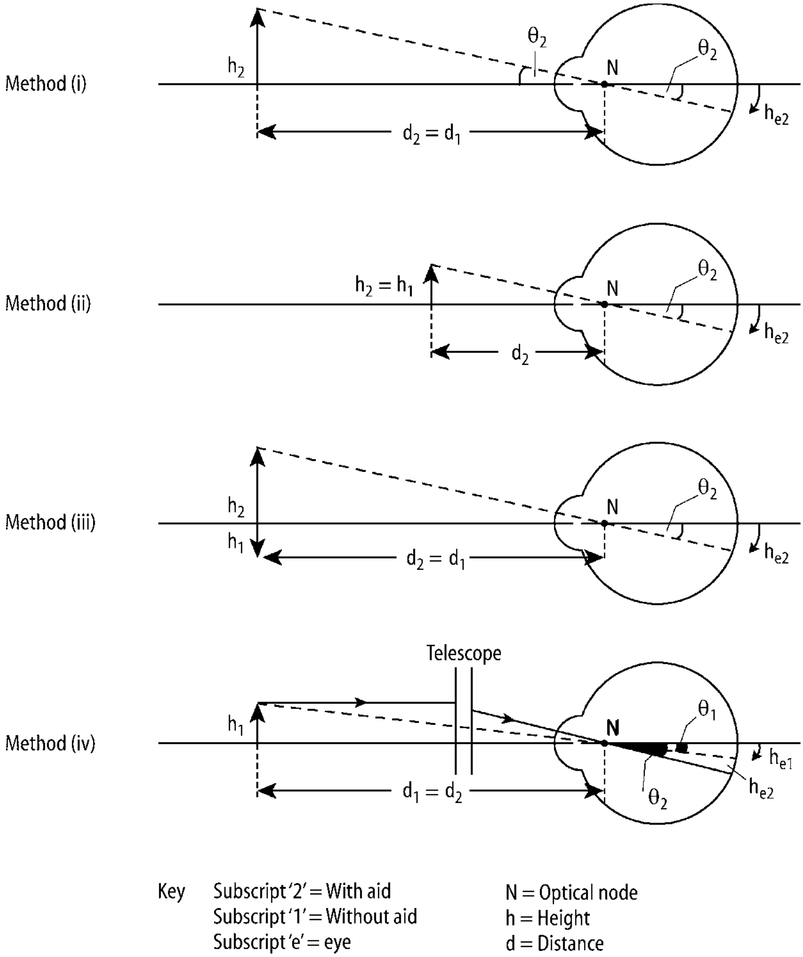


Figure 12.3. Optical geometries for low vision aid methods

4. Telescopic magnification, which makes an object appear larger, without changing the object size or viewing distance

The optical geometries for the four different methods are illustrated in Figure 12.3.

Method (i) Increasing the object size while keeping the viewing distance constant

This method is illustrated in Figure 12.3 (Method (i)), which shows the retention of the same viewing distance (implies $d_2 = d_1$), and the increase in object size, giving $h_2 > h_1$. Therefore,

$$\text{Magnification, } M = \frac{h_2 d_1}{h_1 d_2} = \frac{h_2\, d_1}{h_1 d_1} = \frac{h_2}{h_1} = \frac{\text{new object size}}{\text{old object size}} > 1 \qquad (12.4)$$

The common low vision aid that uses this method is large print, such as large print books, large print newspapers and a range of other large print documents, including bills, accounts, terms and conditions and official letters.

The magnification that can be achieved in large print books is limited by upper limits on the size and weight of a book before it becomes unwieldy rather than technical factors. Typically magnification of 2.5× can be achieved before the book becomes excessively large. However, restrictions of this type do not affect short documents, such as official letters.

Method (ii) Decreasing the viewing distance while keeping the object size constant

This is shown in Figure 12.3 (Method (ii)), where the unchanged object size implies that $h_2 = h_1$, whereas the reduced viewing distance gives $d_2 < d_1$, hence,

$$\text{Magnification, } M = \frac{h_2 d_1}{h_1 d_2} = \frac{h_1 d_1}{h_1 d_2} = \frac{d_1}{d_2} = \frac{\text{old viewing distance}}{\text{new viewing distance}} > 1 \qquad (12.5)$$

This method is most commonly implemented by plus-lens magnification. This gives the user an increased retinal image size by bringing an object closer to the eye without requiring the accommodation normally required to view close objects. The magnification formula for a plus lens assistive system is obtained by assuming that the user has a reading distance of 25 cm or 0.25 m. Since the power of a lens in dioptres is equal to the reciprocal of its focal length in metres, 0.25 m is the focal length of a lens with power 1/0.25 = 4 D. Therefore, the old viewing distance is the focal length of a lens of power 4 D. The new viewing distance is the focal length of the magnifier lens. Therefore substituting these facts in the magnification formula gives

$$\begin{aligned}\text{Magnification, } M &= \frac{\text{old viewing distance}}{\text{new viewing distance}} \\ &= \frac{\text{focal length of 4D lens}}{\text{focal length of magnifier lens}} = \frac{\text{power of magnifier lens}}{4}\end{aligned} \qquad (12.6)$$

where this expression has been obtained by using the fact that the power of a lens is the reciprocal of its focal length.

In practice, the user's reading distance may not be approximately equal to 25 cm. Since the magnification is proportional to the reading distance, users with greater reading distances will obtain higher magnifications and users with shorter reading distances will obtain less magnification. Therefore, users who already hold reading matter very close to their eyes would require a lens with exceedingly short focal length and very high power to obtain any benefit from this approach. Consequently, this method is of most benefit to users who have moderate or high reading distances.

Method (iii) Creating an enlarged image at a constant viewing distance

This method is similar to Method (i). The difference is that an enlarged image rather than an enlarged *object* is produced at the constant viewing distance, as illustrated in Figure 12.3 (Method (iii)). It is therefore sometimes called real image magnification or transverse magnification.

Since the viewing distance is constant, $d_2 = d_1$, while the increase in object/image size gives $h_2 > h_1$, so that

$$\text{Magnification} = M = \frac{h_2 d_1}{h_1 d_2} = \frac{h_2 d_1}{h_1 d_1} = \frac{h_2}{h_1} = \frac{\text{size of image with aid}}{\text{size of object without aid}} > 1 \tag{12.7}$$

Assistive technology systems such as closed circuit television (CCTV) and video-viewers implement this low vision aid principle. Direct measurement of the object and image size can be used to compute the magnification achieved.

Method (iv) Telescopic magnification, which makes an object appear larger, without changing the object size or viewing distance

Telescopic magnification involves the use of a telescope to increase the angle subtended by the image at the eye compared to the image seen unaided. Consequently, the calculation of the magnification requires the exact formula (Equation 12.1) rather than the approximation for small angles (Equation 12.3) used with the other methods:

$$\text{Magnification, } M = \frac{\theta_2}{\theta_1} = \frac{\text{angle subtended at eye by telescope image}}{\text{angle subtended at eye by object}} \tag{12.8}$$

This approach is potentially very versatile, as it does not require any change in the object size or viewing distance. However, there are practical difficulties in developing suitable optical systems and therefore it is not as widely used as it might be. The telescopic principle is used in low vision optical aids for "spotting" telescopes and more rarely as "mobility telescopes".

This theoretical framework of four categories of magnification allows the construction of a very useful taxonomy of low vision aids as shown in Figure 12.4.

Combining the different methods

Overall magnification can be increased by combining the different types of magnification, with the total magnification obtained as the product of the individual factors. For instance consider a person using a CCTV to magnify print of size 6 mm to 48 mm on the screen. This is real image magnification with a magnification factor of $48/6 = 8\times$. If the print is at a distance of 40 cm and the screen is at the reduced distance of 20 cm, this gives a further magnification due to the reduced object distance of $40/20 = 2\times$. Therefore the total magnification is $8 \times 2 = 16\times$.

12.3 Low Vision Assistive Technology Systems

A number of low vision assistive technology systems are described in more detail in this section.

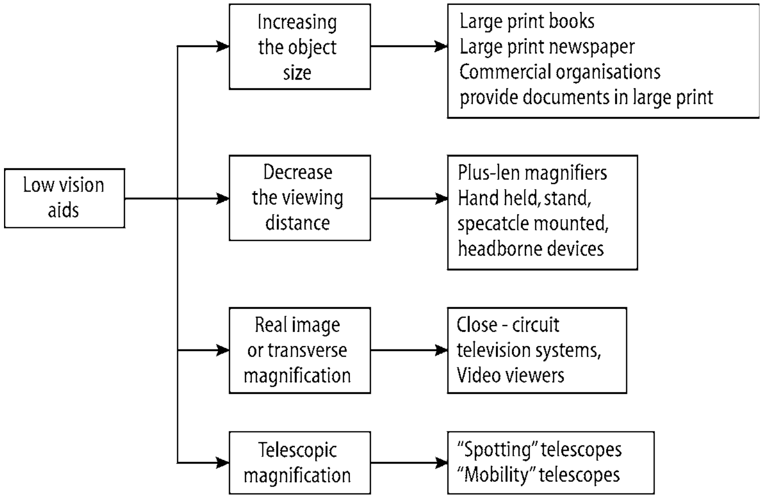


Figure 12.4. Low vision assistive technology systems based on Dickinson's taxonomy

12.3.1 Large Print

Producing large print documents does not just involve making print larger, but also requires consideration of a range of design issues relating to different document parameters in order to make documents accessible to a wide as possible a range of visually impaired people. Government bodies and a number of other organisations have developed guidelines for the production of large print documents. However, it should be noted that different visually impaired people may have different requirements and preferences and therefore the guidelines may involve compromises. Some of the guidelines in Gregory (1996) are presented below (with additional comments by the authors shown in italics) as an illustrative example:

- *Typeface*. Clear typefaces should always be used and unusual, indistinct, or ornate typefaces avoided.
- *Type size*. A minimum size of 14 pt is recommended for people with a visual impairment. The actual size of the typeface should be considered. (Some faces appear larger than others at the same point size). The length of the text should be reduced if necessary rather than the type size.
- *Type style*. Italics and excessive use of capital letters should be avoided, as these letter forms affect the outline shape of words and are therefore more difficult to read.
- *Line length*. Between 50 and 65 characters, inclusive of spaces, *per* line, should be used. Text set in columns may have less than 50 characters *per* line. *However, a large number of fairly narrow columns can also be difficult to read.*
- *Justification*. Ranged left type with a ragged right-hand margin is generally easier to read than justified type. *However, people with some visual processing*

or cognitive impairments may find a ragged right-hand margin distracting and therefore find it easier to read justified text.

- *Layout*. The layout should be uncluttered and logical. There should be contents lists and clear headings at frequent intervals to act as signposts for the reader. The text should be divided into short paragraphs with adequate spaces between and around them.
- *Contrast and colour*. Contrast between backgrounds and type is crucial for legibility. Black on white or light yellow is best for many people. *However, people with dyslexia may require other colour combinations. Some visually impaired people prefer white or another light colour on black or another very dark colour, whereas other people find that this colour combination causes glare.* Note: 1 in 200 women and 1 in 12 men cannot distinguish between red and green.
- *Paper*. Matt paper should be used. Glossy surfaces can cause glare. Thin or semi-transparent papers should not be used, as they allow print to be read through from the other side or other pages. This can interfere with the text being read.

This sample of the available guidelines illustrates some of the issues to be considered in creating large print documents, as well as the possible contradictions between the requirements of different groups of users. If a large print document is being produced for a particular audience, it is advisable to check whether this audience has any particular requirements. Otherwise, the aim should be to create large print documents that are accessible to as many people as possible, while recognising that the result may not fully satisfy all potential users' needs. It is also advisable to consult professional guidelines before producing large print documents (Gregory 1996). However, it should be noted that, though due care and thought is required, producing large print documents is not excessively difficult and a variety of different types of organisations manage to do it relatively successfully.

12.3.2 Closed Circuit Television Systems

Closed circuit television (CCTV) systems are based on the real image or transverse magnification principle. In this case, the user views the magnified image at the normal viewing distance and there is no requirement for any additional optical correction. The usual configuration is a viewing television camera pointing directly or indirectly at the document to be viewed, with the user viewing the magnified image on a television screen. The magnification formula is given by

$$\text{Magnification, } M = \frac{\text{linear size of image on screen}}{\text{linear size of actual object}} \tag{12.9}$$

Most CCTV systems allow the amount of magnification to be adjusted. A maximum magnification of up to 70× can be achieved without optical distortion.

Although the basic formulation for CCTV assistive technology systems dates back to 1959, it was not until the early 1970s that commercial models were produced (Potts *et al.* 1959; Genensky 1969; Genensky *et al.* 1973). Two typical

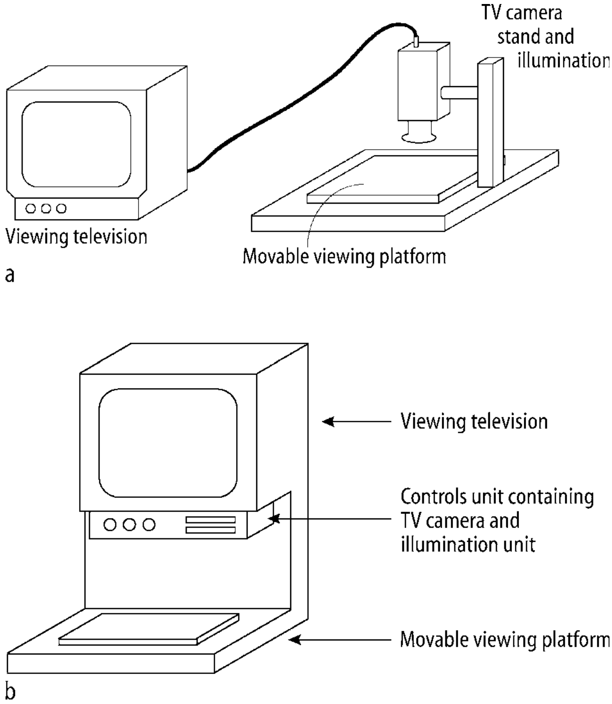


Figure 12.5a,b. Two typical CCTV systems: **a** distributed CCTV system; **b** integral CCTV system

configurations are shown in Figure 12.5. The integral unit has the advantage of compactness. Portable CCTV units are also commercially available. Monochrome monitors are sufficient for text magnification, whereas more complicated systems with colour cameras and monitors are generally required for artwork, diagrams and maps.

CCTVs have a number of advantages and disadvantages (Dickinson 1998). Advantages include the following:

- The use of a zoom lens allows the magnification to be changed rapidly without altering the focus. For instance, low magnification could be used to obtain an overview, before zooming in on a particular section of the text or image.
- CCTVs can often be used for extended periods compared to other optical aids. However, the maximum reading speeds are the same.
- The screen can be viewed with both eyes from the user's standard working distance.
- Many people with a greatly reduced field can read more efficiently with a CCTV than with another optical aid providing the same magnification. This is probably because they can fixate on a particular area of the screen and use the X-Y platform to scan the image through this area. This technique is more difficult to adapt to other optical aids.

- Reversing the polarity of the screen image can be used to give the option of viewing white text on a black background in addition to the standard black on white. This is useful for people with opacities of the optical media, for instance due to cataract, and many CCTV users prefer this mode of viewing.
- CCTV provides greater contrast than optical systems.
- Some CCTVs have the facility for a split screen presentation, allowing different tasks to be viewed simultaneously, for instance in taking notes from a blackboard.

However, CCTVs also have the following disadvantages:

- Acquiring proficiency in CCTV use requires more practice than an optical aid.
- CCTVs are expensive and require regular maintenance and servicing. They are also bulky and, other than miniaturized systems, are not easy to transport.
- Persistence of the image can lead to blurring, particularly with white on black images and limit the maximum reading speed.

12.3.3 Video Magnifiers

The video magnifier is another implementation of the real image or transverse magnification principle. A typical product, the Pocketviewer, is shown in Figure 12.6a and, as can be seen, it is a hand-held video magnifier that is held over the text to be viewed. The device shown in the figure is a colour viewer but there is also a monochrome version available. A magnified image is displayed on the built-in flat screen. A typical unit comprises a 10-cm flat panel housed in a unit of dimensions $3.5 \times 8.6 \times 14.25$ cm. A miniature video camera captures the text on the page and transfers the image to the screen with $7\times$ magnification. The Pocketviewer uses a re-chargeable 1.5 V battery source or can be powered by an a.c. adaptor.

Figure 12.6b shows an internal view of the Pocketviewer that reveals the advanced miniaturisation used in this product. As shown in this photograph, the device has a flat outer case that opens up into two parts with circuit boards in each part. The lid or top part of the case holds a single circuit board that drives the flat screen assembly. (This circuit board is the top one with a small square label in its bottom right hand corner.) The bottom part of the Pocketviewer case contains three components. On the left is the raised camera board with the video camera underneath this board. This small board is easily found since it has the 5 VDC label on the top. The main board in the remainder of the lower part of the case does the signal processing, control, power supply and drives the camera and display tasks. The power supply is on the far right hand side of the main board and in the bottom right corner, the location of the power adaptor socket can be seen.

Two versions are available: the mono version with high contrast black-on-white or white-on-black and a full colour version. The mono-pocketviewer is considered appropriate for reading labels, price tags, credit card slips, restaurant menus, programmes, timetables and similar tasks, as well as browsing through magazines,

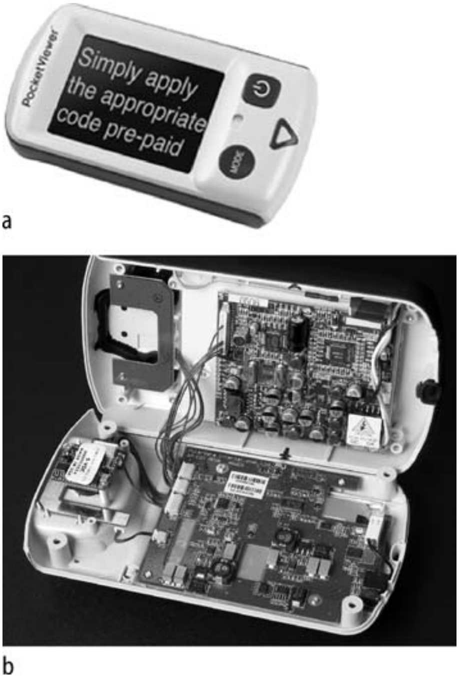


Figure 12.6a. Pocketviewer. **b** Internal view of PocketViewer construction (photographs reproduced by kind courtesy of Humanware Group)

books or newspapers. The colour version will generally be required for maps, photographs, illustrations and three-dimensional objects. It can also carry out all the tasks that can be performed by the mono version.

12.3.4 Telescopic Assistive Systems

Telescopic assistive systems give users an enlarged retinal image while remaining at their normal viewing distance. They can be used for near (writing, handicrafts), intermediate (TV, music, playing cards) and distant (blackboard, street signs, bus numbers) tasks, but have a very restricted field of view. There are special designs for when the user is mobile, but they have the disadvantage of magnifying the object's apparent speed as well as its size. The optical principles of telescopic assistive systems are afocal, with parallel rays of light entering the telescope from a distant object and parallel rays of light leaving the telescope to form an image at infinity.

There are two main types of low vision telescope:

- *Astronomical or Keplerian Telescope* (shown in Figure 12.7a). The image is inverted, as rays from the bottom of the object form the top of the image. The image must then be inverted, so it is the right way round. One possibility is the use of two right-angled (Porro) prisms, as shown in Figure 12.8. The use of

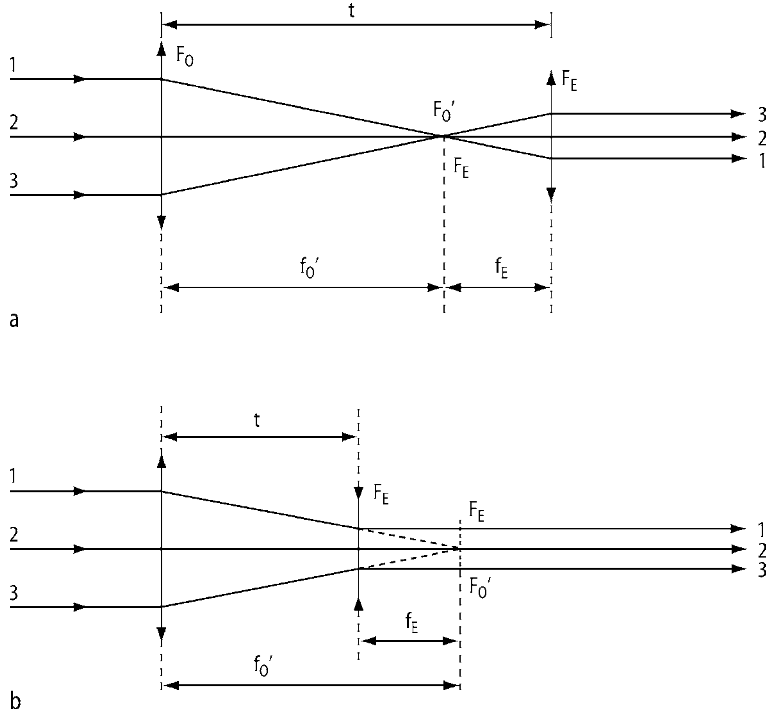


Figure 12.7a,b. Optical geometries for astronomical and Galilean telescopes: **a** astronomical telescope; **b** Galilean telescope

prisms also allows the optical path length between the objective and eyepiece to be 'folded', thereby reducing the overall length of the telescope. The objective lens F_O is convex and forms an image of the distant object at its second focal point. Diverging light from this point is then refracted by the convergent eyepiece lens, F_E. The lenses are positioned so that the first focal point of the eyepiece lens is coincident with the second focal point of the objective lens, so that parallel light will emerge from the telescope system.

- *Galilean Telescope* (shown in Figure 12.7b). The eyepiece lens is positioned so that its first focal point is coincident with the second focal point of the objective lens. Rays of light converging to the second focal point of the objective lens are intercepted by the eyepiece lens and emerge parallel from the system. The image is the right way up, as rays of light from the top of the object go to the top of the image.

The total length, t, of each telescope is equal to the separation of the objective and eyepiece lenses, that is:

$$t = f_O' + f_E \tag{12.10}$$

where f_E and f'_O are respectively the first and second focal lengths of the eyepiece and objective lenses. Using the fact that the focal length of the eyepiece lens is negative and assuming that it is thin, it can then be shown that

$$t_a = f'_{a.O} + f'_{a.E} \tag{12.11a}$$

$$t_g = f'_{g.O} + f'_{g.E} \tag{12.11b}$$

where the subscripts 'a' and 'g' refer to the astronomical and Galilean telescopes respectively and f_E' is the second focal length of the eyepiece lens. It is negative for the Galilean telescope and positive for the astronomical telescope and the objective focal lengths are both positive. Therefore, a Galilean telescope will be shorter than an astronomical telescope with the same magnification, though the difference can be reduced by using prisms to fold the light-path as described above and shown in Figure 12.8.

It can be shown (Dickinson 1998) that the magnification for both types of telescope is give by the negative ratio of the powers of the eyepiece and objective lenses, *i.e.*

$$M = \frac{-F_E}{F_O} \tag{12.12}$$

Since both powers are negative for the astronomical telescope, the magnification is negative for this telescope. This is consistent with the fact that the image is inverted. Since the eyepiece power is negative and the objective power is positive for the Galilean telescope, the magnification is positive for this telescope. This is consistent with the fact that the image is erect.

One of the disadvantages of a telescope is its very restricted field of view. Consequently, the user can only see part of the task at a time and needs to scan across the area to obtain a complete view. The maximum field of view is obtained by making the objective lens as large as possible, whereas the eyepiece lens can be smaller, as long as it does not cut off peripheral exiting rays.

People with myopia or hyperopia (without astigmatism) can compensate by changing the telescope length, adding their full lens prescription to the eyepiece lens or adding a partial lens prescription to the objective lens. Adding the full prescription to the eyepiece lens can be achieved by holding or clipping the telescope

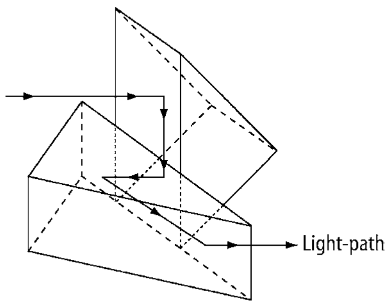


Figure 12.8. Using two right-angled (Porro) prisms to accomplish image inversion and light-path folding

to the prescription glasses or attaching a small auxiliary lens behind the eyepiece lens. Modifying the objective lens is never used in practice. People with myopia would need to shorten the telescope and people with hyperopia to extend it. This is a useful strategy for people with low to moderate myopia or hyperopia and is generally not affected by up to 2 D of astigmatism.

12.3.4.1 Using Telescopic Systems to View Near and Intermediate Distance Objects

Analogous principles can be used to allow near or intermediate distance objects to be viewed. In this case, the full correction for the viewing distance is added to the objective lens or an increased correction is added to the viewing distance to the eyepiece. The telescope can also be focused by changing its length. Changing the length of the telescope by increasing the separation of the objective and eyepiece is often used to adapt a telescope for finite (near or intermediate) working distances. The only restriction is the need to have a tube length of a reasonable size. Astronomical telescopes generally allow a greater range of focus than Galilean ones.

Focusing for near or intermediate distances by changing the telescope length can be considered equivalent to 'borrowing' some of the objective lens power. Consider Galilean and astronomical telescopes with objective and eyepiece powers of 15 D and 45 D (−45 D for the Galilean telescope) respectively. They both have magnification of 45/15 = 3×, namely, +3× for the Galilean telescope and −3× for the astronomical telescope. Using Equation 12.11 and the fact that the focal length in metres is the reciprocal of the power in dioptres gives

$$t_a = f'_{a.O} + f'_{a.E} = 1/15 + 1/45 = 0.067 + 0.022 = 0.089\,\text{m or } 8.9\,\text{cm}$$
$$t_g = f'_{g.O} + f'_{g.E} = 1/15 - 1/45 = 0.067 - 0.022 = 0.045\,\text{m or } 4.5\,\text{cm}$$

When the telescopes are focused on an object at a distance of 50 cm (0.5 m), the vergence of the light reaching the objective lens is 1/0.5 = 2 D. This requires a power of −2 D to neutralise it, which could be notionally 'borrowed' from the objective lens. Then the power of the objective lens of both telescopes becomes equal to 15 − 2 = 13 D. The magnification is now 45/13 = 3.46×, namely, +3.46× for the Galilean telescope and −3.46× for the astronomical telescope. Using Equation 12.11, the lengths of the two telescopes are now increased to

$$t_a = 1/13 + 1/45 = 0.077 + 0.022 = 0.099\,\text{m or } 9.9\,\text{cm}$$
$$t_g = 1/13 - 1/45 = 0.077 - 0.022 = 0.055\,\text{m or } 5.5\,\text{cm}$$

Therefore, in both cases the increase in length is 1 cm.

More generally, if an object is viewed at x m, then the power of the objective lens needs to be reduced by $(-1/x)$ D. Therefore, using Equation 12.12, the magnification is increased to

$$M = \frac{-F_E}{F_O - 1/x} \tag{12.13}$$

As x increases, $1/x$ tends to zero, giving the formula at Equation 12.12 for the magnification of an object at 'infinity'. For small values of x, the reduction in the power of the objective lens and, consequently, the increase in magnification are significant. Since the focal length, f_O of the objective lens is equal to the inverse of the power F_O, Equation 12.13 shows that an object cannot be seen with the telescope at a distance closer than the focal length of its objective lens.

The lengths of the telescopes are now given by

$$t_a = f''_{g.O} + f'_{a.E} \tag{12.14a}$$

$$t_g = f''_{g.O} + f'_{g.E} \tag{12.14b}$$

where

$$f''_O = \frac{1}{F_O + 1/x} \tag{12.14c}$$

A telescope can also be modified to view near objectives by adding a plus-lens reading cap to make a telemicroscope. The magnification obtained is then the product of the magnification of the telescope and that of the reading cap, so that using Equations 12.6 and 12.12 the total magnification is given by

$$M_{\text{total}} = M_{\text{telescope}} \times M_{\text{reading cap}} \tag{12.15}$$

$$= \frac{-F_{T.E}}{F_{T.O}} \frac{F_{RC}}{4} \tag{12.16}$$

where the subscripts 'T' and 'RC' refer to the telescope and reading cap respectively.

12.3.4.2 Telescopes for Mobility

The telescopes described in the above sections are used for sedentary or 'spotting' tasks, as they have a very limited field of view. However, a telescopic device that could be mounted in front of the eye and used to magnify at long distances would be useful for mobility. The approaches used include bioptic and contact lens telescopes.

In bioptic telescopes, the user's prescription lens is a carrier lens with the compact telescope mounted in its upper part. This allows the user to look through the prescription lens as usual and obtain a magnified view of distant objects by lowering their head. There are both lower powered (about 2×) devices, as well as systems with a wider field of view (Bailey 1982). Careful fitting and intensive and structured training are required for successful and safe use (Feinbloom 1977; Kelleher 1979). However, although bioptic devices are used to support mobility and some states in the USA allow driving using them, they are intended for occasional rather than constant use.

A contact lens telescope uses a negative contact lens as the eyepiece and a positive spectacle lens as the objective of a Galilean telescope (Bettman and McNair 1939). The length of the telescope is equal to the separation of the lenses. This is the only type of magnifying system that would probably be allowed to be used for driving in the UK. Contact lens telescopes have limited magnification and they have been found to be particularly useful to people with congenital nystagmus, who often show a much greater than expected improvement in visual acuity.

12.4 Audio-transcription of Printed Information

Blind and visually impaired people require access to a wide variety of different types of printed information, including books, newspapers, menus and timetables. One of the earliest approaches to making print accessible was the talking book. This involved making a recording of the book being read, generally by volunteer readers. Once the original recording was made, multiple copies were produced, originally on tape and currently on cassette or CD. The copies were then distributed to be played on an appropriate player. This approach has also been used to produce talking newspapers. The main advantage of this approach is that the recording sounds natural since a person, rather than synthetic speech has been used to produce it. The drawback of this is that it is time intensive and expensive if the recordings are made by paid staff. In addition, it is most suitable for items such as books with a stable text that will be used for an extended period. It is less practicable for items such as menus, timetables and theatre programmes that are constantly changing.

This gives the need for reading systems or devices that can read items as they are presented, rather than having a recording prepared in advance. Such reading systems generally include text-to-speech conversion software, which is discussed in more detail in Chapter 14. A simple classification of reading systems is given in Figure 12.9. It should be noted that one of the main distinctions is between stand-alone reading systems and reading systems that are computer-based. Stand-alone reading systems and the Read-IT project are discussed in the next two sections.

DAISY technology (discussed in Chapter 15) has been developed as a standard for audio output of printed material. The idea of a standard navigable format for visually impaired end-users to access information in audio format is clearly a good one. However, the time lag due to number of factors including the time spent in developing the format and working for acceptance of it, have meant that the technology has evolved. Despite considerable hard work to publicise DAISY, it has not been taken up on a large scale by publishers and many publishers are unaware of it or how they could use it. It is hoped that this situation will change in due course.

From the end-user perspective there are advantages in having output which can be played on widely available standard devices, such as a CD, cassette or MP3 player, rather than requiring a special player. As this indicates, there are advantages in a design for all approach to providing information in audio format to anyone and everyone who might want to use it, including visually impaired and other print disabled people. This may mean revisiting and updating the DAISY standard from a design for all perspective or ensuring that visually impaired and other end-users have a choice of a number of different formats, including DAISY, so they can choose the one that is best suited to their needs or even use different formats in different circumstances.

12.4.1 Stand-alone Reading Systems

Stand-alone reading systems are independent systems that are able to able to scan a printed document (including letters, books, leaflets and newspapers) and

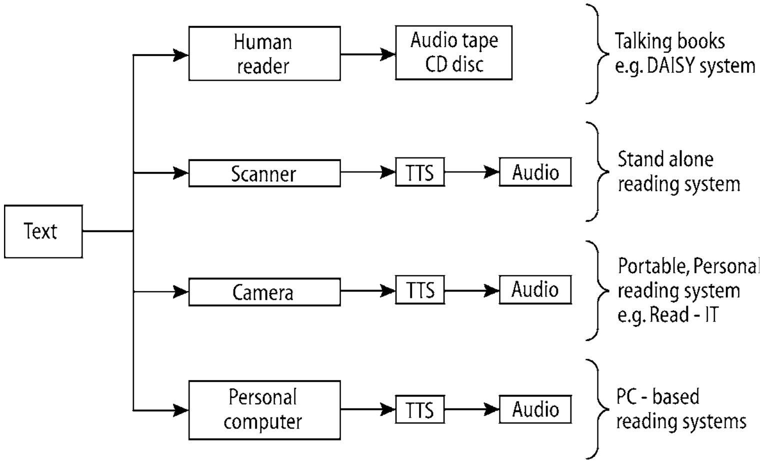


Figure 12.9. Stand-alone text-to-speech (TTS) technologies

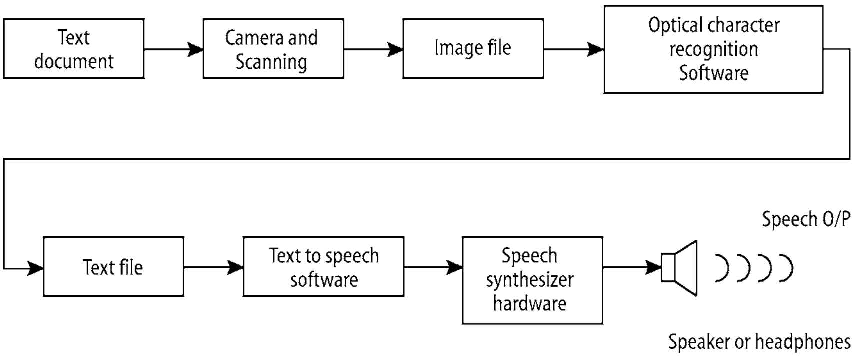


Figure 12.10. Block diagram of reading system operations

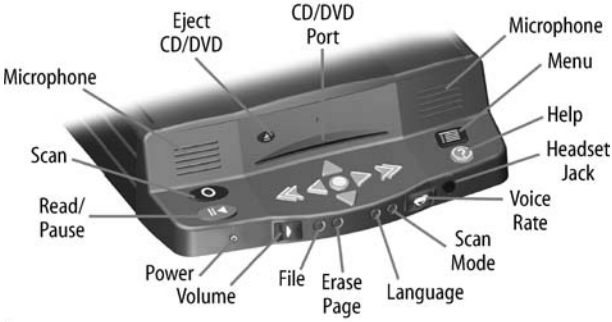
produce an audio (or tactile) version of the document for visually impaired and blind readers. The sequence of operations carried out by reading systems is shown in Figure 12.10.

These operations comprise the following three main stages:

Stage 1. The camera and scanning mechanism create an image file.
Stage 2. Optical recognition software and/or hardware converts the image file to a text file.
Stage 3. Text-to-speech software uses the text file to drive a speech synthesizer card and speaker unit, thereby producing an audio speech output.

Commercial reading systems generally have alternative input and output options. For instance, on the input side the system may be able to read from CD ROM or from stored files and, in addition to audio output using a speaker and headphones, the output may be displayed as text on a computer screen.

a



b

Figure 12.11a,b. Scanning and Reading Appliance, SARA™: **a** SARA™, in action; **b** SARA™, the control panel (photographs reproduced by kind permission of Freedom Scientific, USA)

Figure 12.11 shows photographs of the Scanning and Reading Appliance (SARA™) developed by Freedom Scientific, USA. Some of the key functions of SARA are listed in Table 12.1. As can be seen from Figure 12.11 and Table 12.1, SARA is a highly convenient route to the audio transcription of printed information. However, it is not very portable and therefore more suitable for applications in a fixed location than for use while moving around. A portable system, called Read It, is discussed in the next section.

12.4.2 Read IT Project

Portable devices generally have the advantage of reducing costs since the same device can be used in different locations and is easier to handle. In the case of reading systems, there is a wide range of textual information, including menus, price tags, bus and train timetables, indicator boards and theatre programmes, found in different locations that could not be read with a fixed system. In addition to the technical issues associated with portability, further technical challenges are posed by the wide variety of material to be read, the difficulties associated with reading hand written and poor quality texts and the fact that some textual information, such as street signs and indicator boards, generally have to be read at a distance.

Table 12.1. Some technical feature of SARA™ reading appliance

Controls	Large, colour coded with tactile markings and symbols See Figure 12.11b for layout Search facilities: single word; single line, fast forward, rewind, move up page, move down page
Speech control	Controls for speech rate, and volume, selection of the voice from a voice set, choice of languages from 17 options
Input	Scanned documents, background scanning operation Files (.txt, .rtf, .doc, .pdf, .html) CD ROM drive, DAISY books, microphone input
Output	Stereo speakers (integral to appliance), audio jack for headphones Text output to computer screen with display options
Some technical specifications	Power: 100–240 V, rear power jack input Size: 50.8 × 8.89 × 30.48 cm Weight: 8.16 kg 20 GB hard disk drive; 256 MB RAM; 600 MHz processor

One approach to producing a portable reading system for blind people is the prototype Read IT project (Chmiel *et al.* 2005) carried out by a student team from the Department of Computer Science and Management of Poznan University of Technology, Poland under the guidance of Dr. Jacek Jelonek.

Enduser aspects of the Read IT system

End-user involvement in the development of (assistive) technology systems from the start is crucial to ensure that the resulting device does meet the needs of the end-user community, and reduce the likelihood of it being rejected. In the Read IT project the development team worked with the Polish Association of Blind People to draw up a list of end-user requirements, which included the following (Chmiel *et al.* 2005):

1. The device should be comfortable (portable and lightweight) to wear and should integrate the user into the wider community not identify them as different.
2. The user should have their hands free to engage in other activities whilst listening to the speech output.
3. The user should be able to hear other sounds as well as the generated speech.
4. The user should be able to move onto other tasks once positioning and capture of the text is complete.
5. The generated speech should be clearly understandable and resemble human speech.

These requirements from the end-user community were translated into design specifications and influenced the final design and implementation. Requirement 1 led to the device being lightweight, portable and as unobtrusive as possible. Requirements 2, 3 and 4 arise from safety considerations and the requirement for the user to, for instance, have their hands free to use a long cane and/or carry shopping while using the device. They resulted in the speech output being delivered to only one earphone that was directly inserted into the ear. Requirement 4 enables the end-user to either relax or move onto other tasks once the text to be read has been located and captured. Requirement 5 has been translated into a design specification for the quality of the speech synthesizer card used in the device.

Engineering issues and implementation

As illustrated in Figure 12.12, an image of the text is captured by a video camera positioned in the user's sunglasses. This image is analysed and the text content identified drives the speech synthesizer card. The resulting speech output is then delivered to the user *via* a single earphone. The signal processing unit, speech synthesizer card and battery power supply are housed in a small box worn at waist level. Manual control is *via* a small hand-held Braille key pad. A full description of the development process is given in the Read IT report (Chmiel *et al.* 2005). In view of the requirement that the device should integrate the user into the wider community, the video camera could presumably be worn on glasses with plain lenses at times when there is little sun. However, difficulties could be encountered in transferring the device between different spectacle frames.

In contrast to technologies, such as mobile phones, where miniaturisation causes difficulties for blind and visually impaired people, it is component miniaturisation that has made portable reading devices, such as Read It, feasible. In particular, an important feature is the miniature video camera that can be unobtrusively mounted on the user's sunglasses. A PVI-430D video was selected and this captures 30 frames per second at a resolution of 640 × 480 pixels. The very small size of the video

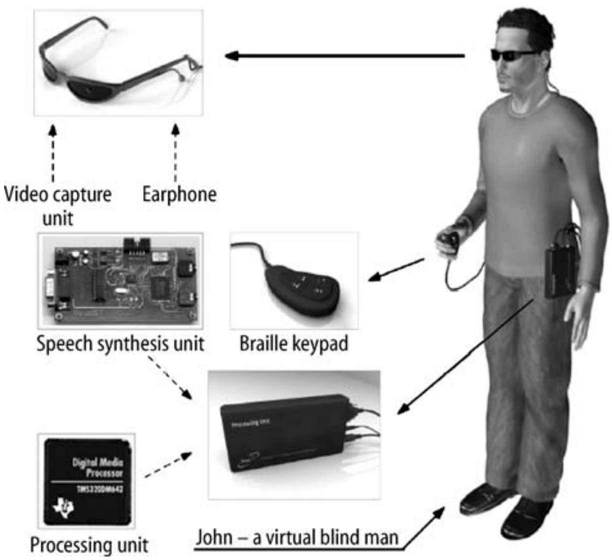


Figure 12.12. Overview of the Read IT system (Chmiel *et al.* 2005)

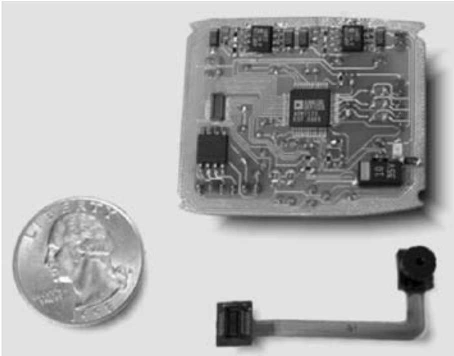


Figure 12.13. Video capture unit (Chmiel *et al.* 2005)

capture unit can be seen from Figure 12.13 where the microcontroller chip is just 5×5 mm. The camera range is between 0.4 and 0.8 m for standard sized fonts. It is generally feasible to approach to within this distance of bus timetables, indicator boards and street signs.

Software issues

The Read IT project used a mix of standard software and customised software and DSP algorithms developed by the project. The steps involved in the digital signal processing algorithms required are shown in Figure 12.14.

Two aspects of this digital signal processing architecture are especially interesting. First, there is a "navigation task" with associated "navigation messages". This module generates voiced directional instructions to the user to ensure that the camera is directed at the text to be read. The audio message feedback loop was designed to optimise and enhance image capture and identification within the device. Once a satisfactory image has been captured, then the important operations of analysing the video text captured can proceed. This involves a number of subtasks, including text segmentation, enhancement and recognition. This is

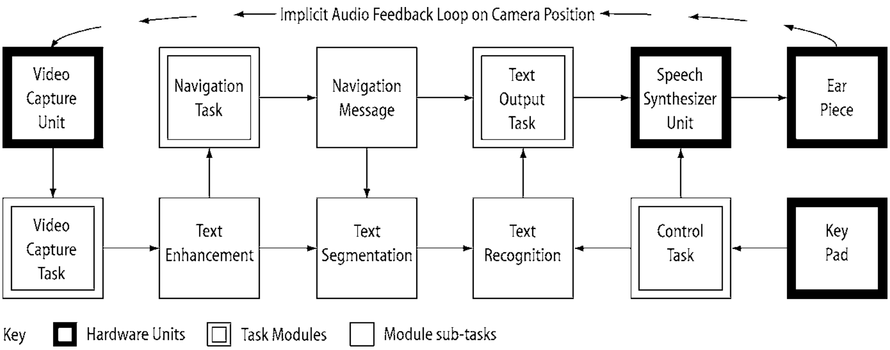


Figure 12.14. Digital signal processing framework for Read IT (Chmiel *et al.* 2005)

the second aspect of the digital signal processing architecture that was important for technological success. The Read IT project involved the development of novel algorithms to accomplish these steps.

Testing and cost

Extensive laboratory testing was carried out to assess the technical performance of the design. Limited end-user assessment has been performed and further tests are planned. The costs of the prototype system are moderate. However, though the prototype seems to have been very successful, there remains the issue of whether the Read IT system can be developed further and either commercialised or delivered to end-users by a social route. The fact that the development team won the second prize at the IEEE CSIDC event in 2005 may be helpful, and the team are intending to progress the prototype further.

12.5 Tactile Access to Information

Tactile access is required to a wide range of different types of information, including textual information, graphical information such as maps and pictures, and symbolic representations such as music and mathematics. A full discussion of the tactile science, haptics, is presented in Chapter 4 in the fundamentals section of this book. Whether tactile or audio access to information is preferred will depend on both the user and the type of information. Some groups of users, such as deafblind people, may require tactile information. In some cases, such as the representation of graphical material, users may require a combination of audio and tactile information. Although Braille is the main tactile method for the representation of print, only a small proportion of blind people read it. Another tactile language, Moon, is used mainly in the UK. The Fishburne alphabet is used mainly for labelling. A more detailed presentation of the technologies used for Braille conversion is presented in Chapter 14 and a discussion of the Braille representation of music scores is given in Chapter 16.

12.5.1 Braille

Braille is a system of reading and writing for blind and deafblind people. It is named after its inventor Louis Braille who developed it from a system called night writing that was developed for communicating with troops after dark, but was not very successful. It uses raised dots, often embossed into heavy paper, to represent the alphanumeric characters and a number of common words. The raised dots are organised into rectangular 6-dot cells, with 64 different configurations, as shown in Figure 12.15. Readers use their fingertips to interpret the tactile character representations and proficient Braille readers develop techniques for doing this efficiently.

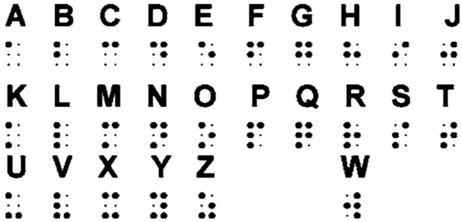


Figure 12.15. Letters in Braille

Each letter is represented by a particular arrangement of six dots. Numbers are represented by a numeral symbol followed by the letters A–J to represent the numbers 1–10. There are also short forms for some common words and combinations of Braille dots in a cell for contractions of two or more letters.

There are two types of Braille: Grade 1, which is a simple letter-by-letter transcription of text, and Grade 2, which is a condensed form that employs the combination forms for common words and letter sequences. Grade 2 Braille is shorter, quicker to read, less bulky in paper use, cheaper to produce and is used by experienced Braille readers. This form of Braille, which extends to 189 abbreviations and contractions, can give a space saving of 25% compared to a Grade 1 Braille book.

About 13,000 people in the UK regularly read Braille out of 1 million people who could be registered as blind or partially sighted. This low proportion is due to the difficulty of learning Braille by people who lose their vision later in life, the fact that some blind and partially sighted people are still able to read text and that some causes of vision impairment such as diabetic retinopathy also affect the sense of touch.

12.5.2 Moon

Moon was developed in 1847 by Dr William Moon. Since it has similarities to ordinary print characters (see Figure 12.16) it is easier to learn by people who have previously read visually. However, it has the significant disadvantage of having about 80 times the volume of the print version and four times that of the Braille version. The high cost of production has meant that very few books are produced in this medium. The number of Moon readers has dwindled to about 400, most of whom are in the U.K. Until recently Moon was produced boustrophedon, which had the advantage of not requiring back tracking from the end of one line to the beginning of the next and the disadvantage of requiring alternate lines to be read in different directions.

12.5.3 Braille Devices

There are a number of Braille devices available commercially and this section provides a brief glossary of the main types of devices.

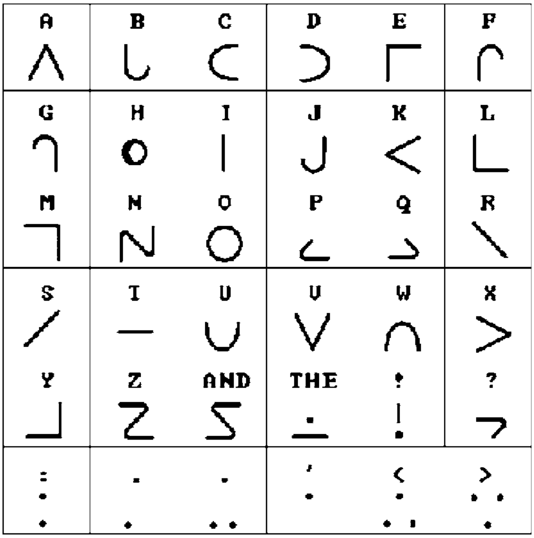


Figure 12.16. Moon alphabet

Brailler. This is a manual machine rather like a heavy-duty typewriter that embosses the Braille cells onto thick paper to give Braille text.

Braille notetaker. This is a Braille version of a personal data assistant. It is a small portable device for recording and retrieving notes, dates, diary, calendar and other personal data. They usually have an integral Braille keyboard, a refreshable Braille display and generally, a speech synthesiser interface. Typical Braille notetakers are shown in Figure 12.17.

Braille keyboard. This is a computer input device that provides Braille input to a computer in the same manner that a QWERTY keyboard is used to input text to a computer.

Braille displays. Refreshable Braille displays or soft Braille displays are output devices for reading text from a computer screen or file in Braille cells.

Text-to-Braille transcription software. This is computer software designed to transcribe text into Braille (and sometimes *vice versa*) in preparation for being queued to a special Braille embossing printer for the production of a hard copy.

Braille embossing printer. This is a printer used to produce embossed Braille paper copy output. These embossing printers can be used to print Braille as output from a computer or print a prepared Braille file.

These Braille computer input and output devices are all discussed in more detail in Section 12.6 on accessible computer systems.

Figure 12.17. Braille notetakers (photograph of the Braille Lite Millennium 20 and 40 by kind courtesy of Freedom Scientific, USA)

12.6 Accessible Computer Systems

The widespread use of computers and information technology make it particularly important that they are fully accessible to blind and visually impaired people. In a recent investigation of deafblind computer users in the U.K., Hersh and Johnson (2005) found that even in a community with a wide variety of *dual* sensory impairments, computer systems could be configured successfully for accessibility. It should not be a surprise that once the accessibility hurdle was overcome, the deafblind computer users often found their computer an indispensable part of daily living. Indeed, one deafblind survey participant commented that without the Internet access to a supermarket ordering and delivery service they would starve!

From the point of view of the user, it is the accessibility of input and output that are important. Computer input is generally *via* the keyboard and mouse, with the keyboard used to enter data and to give some commands to the system and the mouse largely to access commands and other options of the menu-based operating system and graphical user interface. Although the keyboard and mouse are themselves tactile devices, feedback that the correct key has been pressed is generally obtained visually from the screen. The output is generally displayed on the screen, possibly supplemented by audio over speakers and can be made more permanent by printing a paper hardcopy and recording the audio output. None the less, computer operating systems are largely based on graphical user interfaces.

Therefore computers and, to a lesser extent computer inputs, are inaccessible to many blind and visually impaired people. Fortunately however, *design for all* principles have had some influence on the design of computer operating systems. Consequently, operating systems such as Windows and Linux have a number of options for customisation to meet the user's specific needs and to increase accessibility to a wide range of users. In Windows, these customisation options are found in both the accessibility and other toolboxes in the Control panel. Newer versions of Windows have an increasing range of options for customisation and accessibility. While this is a positive and encouraging trend, it has the disadvantage that visually impaired and blind users with older software will not have access to many of these options.

Thus, computer accessibility concerns the further aspect of configuration in addition to input and output. Another dimension of accessibility relates to the sensory modalities used to access the system. Since computer systems have not yet been developed which use taste or smell, this gives the three sensory modalities of sight, hearing and touch for accessing computer information, *i.e.* the provision of information in visual, audio and tactile form. The matrix of options for accessible computer architectures is presented in Figure 12.18.

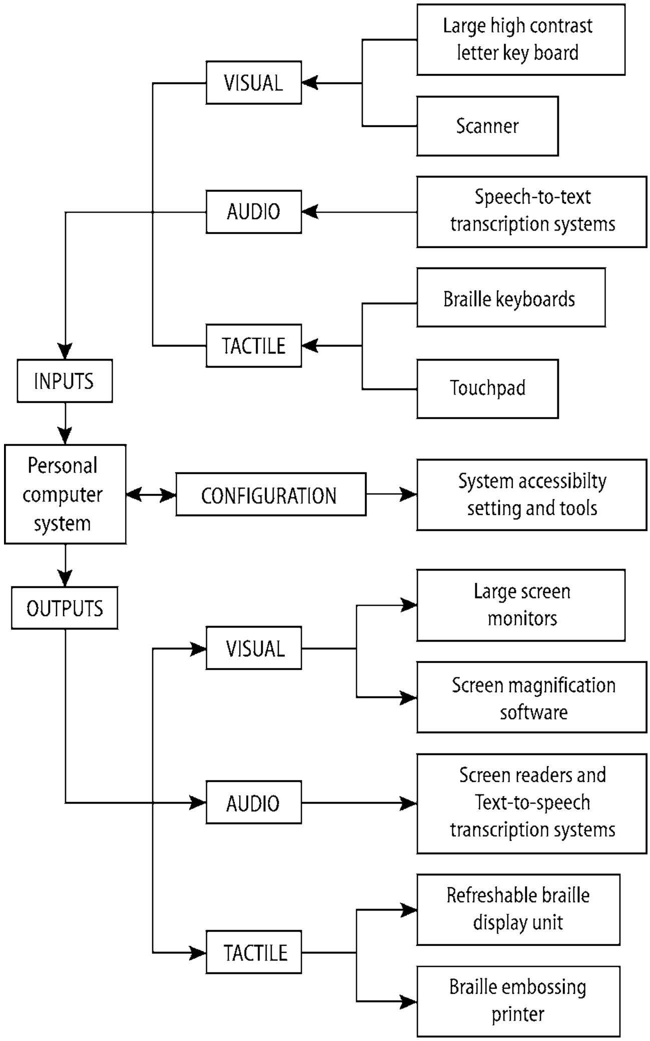


Figure 12.18. Generalised accessibility architecture for PCs

The system configuration options available with both Windows XP and Linux will be considered briefly here. It should be noted that they are similar, but not identical. The options available with Windows XP include those listed immediately below. Other operating systems will provide slightly different configuration options and may use different labels for some of these options:

- *Mouse.* The appearance, size and click speed of the mouse pointer can be selected to make it easier to see and give a reduced click speed for users with reduced manual dexterity or slower reactions. The mouse can also be configured for right or left hand use. A click lock function permits the mouse to be used to highlight and/or drag objects without holding down the mouse button. There is an option for indicating the pointer location by concentric circular ripples when the 'Ctrl' function key is pressed. People with tunnel vision may find this accessibility option particularly useful. There is also an option to move the mouse automatically to the default option in a check box.
- *Sounds.* Users can configure the system to select the operations for which they do and do not want audible alerts. They can also choose the type of sound from a number of options to ensure that the alert is effective and to identify what the alert is for. There is also the option to turn off all sounds, which is useful to noise-sensitive users. There are also a number of different options for controlling the volume settings for both sounds and voice output. 'SoundSentry' can be used to give a visible alert whenever the system makes a sound and 'ShowSounds' can be used to display captions for speech and sounds.
- *Keystroke options.* 'Stickykeys' can be used to combine 'shift', 'Ctrl', 'Alt' or the windows logo key with another key by pressing the keys in sequence rather than simultaneously. 'Filterkeys' can be used to make Windows ignore brief or repeated key strokes or reduce the repeat rate. 'ToggleKeys' can be used to give an audible alert when pressing the capital, number or scroll lock keys.
- *Display.* High contrast with a number of colour and size options can be used to improve the readability of the menus and other information provided on the screen. The blink rate and width of the cursor can be varied. In particular, cursor blinking can be turned off, which is particularly useful for people who find this irritating or stressful.

Linux (LINUX, undated) is a free Unix-type operating system, which was developed under the GNU General Public License (GNU, undated). As required by this license, the source code for Linux is freely available to everyone. Its graphical user interface can be configured in several different ways, with two of the main options called KDE and GNOME. The system customisation options available with Linux GNOME will be considered here and include the following:

- *Accessible login.* Enables users to log in to the desktop even if they cannot easily use the screen, mouse or keyboard in the standard way. This feature also launches assistive technology on log in and gives the login dialogue the desired visual appearance, for instance high contrast.

- *Keyboard.* The keyboard preference tools give a number of options, including configuring the keyboard to emulate the mouse. In mouse emulation the numeric keypad is mapped to mouse functions, with '/', '*' and '-' representing the three mouse buttons, '5' and '0' the button click and toggle respectively and the remaining numbers moving the mouse pointer round the screen. The toggle mouse button enables the currently selected mouse button until the mouse button click key is pressed, allowing drag and drop operations to be carried out. Slow keys can be used to specify the duration a key must be pressed and held before it is registered and to give an audible indication when a key is pressed, accepted or rejected. Bounce keys enable rapid repeated keypresses to be ignored and repeat keys allow long keypresses of the same key to be ignored and control the repeat rate of a keypress. Sticky keys allow key combinations to be pressed in sequence rather than simultaneously. In the latch mode, the modifier key (alt, ctrl or tab) is pressed once and remains active until another key is pressed. In lock mode, the modifier key is pressed twice and remains active until it is unlocked by being pressed a third time. The keyboard can also be customised to give an audible notification when a toggle key, such as Num Lock or Caps Lock, is pressed to indicate whether it is being set or unset.
- *Desktop.* Themes can be used to customise the appearance of the desktop or different components of the desktop can be customised individually. The Theme preference tool can be used to give a desktop with high contrast colours (either dark text on a light background or light text on a dark background) and/or large print. The Font preference tool can be used to change the font used in the desktop background and desktop applications. The icon settings associated with a theme can be modified to be high contrast, either dark on light or light on dark. In addition, the user can create their own themes.
- *On-screen keyboard.* This application displays a virtual keyboard, which can be operated using the standard mouse pointer or an alternative pointing device, on the desktop. There are two types of virtual keyboard, one contains alphanumeric characters which can be used to compose text and the other contains keys that represent the applications running on the desktop or the menus in these applications.
- *Screen reader and magnifier.* The screen reader gives access to standard applications on the desktop using speech and Braille output. The magnifier provides automated focus tracking and full screen magnification.
- *Mouse.* The period of time between the two clicks of a double click can be increased, the size of the mouse pointer can be changed and the mouse can be configured for left-handed-use. Although the default version of GNOME does not include mouse pointer themes, these can be installed to further change the mouse appearance. The speed at which the mouse pointer moves around the screen when the mouse is moved and the sensitivity of the mouse pointer to movements of the mouse can be configured. High settings require only small movements of the mouse to cover the whole screen, whereas low settings require the mouse to be moved larger distances. The threshold distance that an item

must be moved before this as interpreted as a drag and drop action can also be selected. The Highlight the pointer option highlights the pointer when the control key is pressed.

- *Cursor blinking*. This can be turned off in text boxes and fields.

The keyboard accessibility status panel indicates which accessibility features are enabled.

12.6.1 Input Devices

12.6.1.1 Large Letter Keyboards

Standard keyboards can be modified by stick-on keytops, which are sets of high visibility letter, number and function keys. Various high contrast colour combinations are available, including black letters on a yellow, white or beige background, blue letters on a white background and bright yellow letters on a black background. The characters are large and bold, as well as both upper and lower case. The key tops sometimes also have Braille embossing.

12.6.1.2 Scanners

A scanner consists of a camera and a scanning mechanism. It is used to create an image of a print document or picture. Optical character recognition (OCR) software is then required to convert the image of the scanned document into a file that can be further processed, as discussed in Chapter 15. The file can then be edited using a word processing package or read using text-to-speech software.

12.6.1.3 Speech-to-text Transcription Software

For a number of years it has been the dream of many computer users, both sighted and visually impaired, to be able to speak instructions to a computer rather than enter them more laboriously through a keyboard. This is not yet completely possible due to the basic computer limitations of interpreting instructions literally. Therefore, it is not possible to, for instance, give a computer an audio summary of the basic contents of a letter or other document and expect it to produce a coherent text. However, it is possible to dictate the text of a document and for speech-to-text transcription software to turn it into text on the computer or to give audio instructions to the computer, for instance to *open*, *copy*, *print*, *save* or *delete* a file.

The quality and accuracy of speech-to-text systems has increased significantly since they were first introduced. The accuracy of speech recognition can also be improved by a period of training for the system in recognising a particular voice. However, accuracy still tends to be best for people with clear voices and standard accents using a restricted vocabulary. Errors due to incorrect recognition of commands by the system could cause serious problems if they occurred. Fortunately, serious errors of this type are unlikely, as the set of commands involves a restricted vocabulary that the speech recognition system can be trained in. In addition, commands with significantly different meanings sound very different. Therefore, it is highly unlikely that the system will delete a file rather than save it. However, there could be confusion between similar sounding commands such as 'save' and 'save as'. Errors are more likely to occur in inputting text, due to the much wider vocabulary and greater variations in language and therefore careful editing will be required. People with unclear speech may experience very unsatisfactory performance.

Speech-to-text recognition is based on pattern recognition. Pattern recognition software has two components: an analyser and a classifier. The analyser produces a spectral representation of the speech signal based on what are called feature vectors. The feature vectors are then input to the classifier which decides on the particular class a given feature vector or vectors belong to. The main difference between numerical and structural classification is that numerical classification assigns each feature vector to a particular class, whereas in structural classification the assignment is of a sequence of feature vectors rather than a single feature vector. These and other aspects of speech-to-text systems are discussed in more detail in Chapter 14.

12.6.1.4 Braille Keyboards

Braille users may prefer to use a Braille keyboard. Fluent Braille users are likely to have little difficulty in learning to touch type, which would give them access to a standard keyboard. However, they will still require some form of audio or tactile feedback to avoid errors. Braille computer input devices are not very readily available, probably due to the fact that the size of the potential market is considered to be too small and decreasing.

The simplest approach to producing a Braille keyboard involves converting a standard keyboard using stick-on transparent Braille embossed keytops. This is the simplest and cheapest method. It has the advantage of providing the user feedback on the keys pressed, but requires touch typing skills or at least approximate memorisation of the keyboard layout or data entry will be very slow.

There are a number of notebook PCs and personal notetakers designed for Braille users. They have integral Braille keyboards, refreshable Braille displays and generally, a speech synthesiser interface. The notetakers are able to exchange files with personal computers, giving the Braille user computer access using the notetaker's Braille keyboard. See Section 12.6.4 for more information on accessible portable computers.

There are also Braille keyboards that can be used directly with desktop and portable PCs. They generally have ten keys, eight for data entry and two keys that are used as space bars as shown in the sketch of Figure 12.19. Unlike the alphanumeric keyboard generally used with a PC, where one keypress is required for each letter or symbol, Braille keyboards have a chordic approach with a number of keys being pressed simultaneously to enter each letter or symbol. Some Braille keyboards have 'hot keys' that allow the user to toggle between Grade 1 Braille and

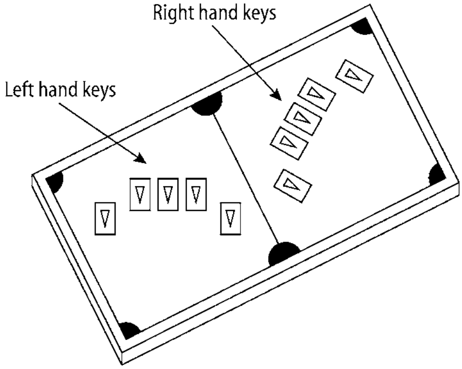


Figure 12.19. Sketch for Braille keyboard

Unified Braille Code. Combinations of the eight keys can be used to type all Latin-based alphabet languages and all computer data, as well as to obtain the function keys, cursor control keys and frequently used key combinations, involving 'shift', 'control', 'alt' and 'delete'. Braille keyboards frequently have an editing mode that is entered by pressing the two space bars simultaneously. In this mode, the user can backspace and delete text or move the cursor to the left or right.

Since only ten keys are required, the keyboard size can be significantly reduced as required for notebooks PCs and notetakers, without making the keys very small, as sometimes happens with standard keyboards on laptop and palmtop computers. Since each finger is assigned to only one key and the only movement required is vertically down to touch the key, the occurrence of errors due to pressing the wrong key or missing a key is reduced. It has been suggested that chordic keyboards have a number of ergonomic benefits with regards to reducing repetitive strain injury. It has also been suggested that chordic keyboards have the fastest learning curves and lowest error rates. There are clearly significant benefits for Braille users, who are not required to learn a new keyboard layout. Other users would have to become accustomed to the relationship between the keys pressed and the resulting letter or symbol. However, this is nothing like as complicated as learning Braille, as it does not require the user to be able to read Braille either tactilely or visually. It only requires them to learn the key combinations equivalent to the different letters, symbols and function keys. A Braille keyboard also provides a potential communication approach for deafblind people, which may be faster than other types of fingerspelling. The Screen Braille Communicator is one portable, two-way communication device designed for use by deafblind people that exploits this potential (Lagarde 2007). Despite suggestions that Braille keyboards make all other types of keyboards and keyboarding technologies obsolete, they are still considered the preserve of the small number of blind and visually impaired people who know Braille.

12.6.2 Output Devices

12.6.2.1 Large Screen Monitors

Computer screen technology has made significant advances in the last decade, with high definition flat screen technology increasingly replacing the previous technology of vacuum-tube-based computer monitors. This change in technology has been highly beneficial to visually impaired and blind computer users, since it has significantly increased the clarity and definition of the screen output. In practice, this has had the same effect as a degree of magnification for a number of visually impaired computer users. The use of large screens is also equivalent to magnification, often without the disadvantage of only being able to magnify a portion of the screen image or only being able to fit a portion of the magnified image on the screen. When the whole task cannot be fitted onto the screen, there is an increased risk of the user losing their place.

Unfortunately, large screens are more expensive than smaller ones and it is small or medium size screens which are available in libraries and other public buildings and which are generally supplied as standard with computer systems. Currently the largest available monitor is about 53 cm measured diagonally across the screen panel. The upper limit on size is probably determined by practical considerations, such as the size of the available desk space and the screen width users can easily scan across, rather than technical factors.

12.6.2.2 Screen Magnifiers

A visually impaired person with residual vision can use a screen magnifier to access computer and information technology. A screen magnifier magnifies the screen using a small window containing an enlarged area of text. Generally, only part of the screen is enlarged at any one time. The window facility may be mobile, in which case it can be moved to a new text area rather like a magnifying glass.

There are several types of software available and these can magnify the screen by up to 20×. The main difficulties with screen magnifier software is teaching the user how it works and how to recognise the working location on the screen. Users of magnifying software may initially experience problems in locating the mouse on the screen, and identifying exactly what text is being shown. Determining the spatial location of other windows and icons on the screen can also be a problem.

There are also hardware magnifiers, including the Super Vista product. This may be an expensive solution for individuals and it may be more appropriate for institutional and commercial use. Hardware magnifiers completely replace the PC's graphics system and can work with any graphics program. Magnification is from 1.5× to 16× full-size, and there are usually a number of different ways of displaying the screen text. Screen magnifiers are discussed fully in Chapter 13.

12.6.2.3 Screen Readers

Screen readers are software products that support print disabled people, including blind and visually impaired and dyslexic people, by reading out the text on the screen and outputting it to the user *via* a speech synthesiser or a Braille display. There are a number of different screen readers available with different features. Some screen readers are combined with screen magnifiers. The different screen-readers vary greatly in price.

As each screen of new information is presented, the screen reader only outputs the part of the screen which is active. It may also scan the screen and make decisions on which information should or can be output to the user. This means less delay for users who are only interested in part of the screen, but requires the system to be configured appropriately so that users receive all the information they require. If information has been generated by a third party, then the format has to be readable. In the case of website pages, there are a number of guidelines for accessible page design (See Section 12.7).

Some screen readers also add information about the layout of the page. Other screen readers simply scan along each line of text and output exactly what is input. If text is embedded within a table then the user will be informed that the information is contained within a table. Some screen readers are able to scan along menu bars and tool bars and choose a particular option, which is determined by the audio or Braille output to the reader. Many screen readers can read back typed input as the user is typing it in. This means that the user can check that their typed input is correct.

Screen readers involve the production of synthetic speech from text. There are two main approaches, parametric and concatenative, with concatenative speech synthesis most commonly used now. Parametric speech synthesis is based on a model of the human articulation system and control of a number of parameters, such as gain, switching between voiced and unvoiced sounds and the frequency of voiced sounds, to produce speech sounds. Concatenative speech synthesis involves the combination of very short segments of speech produced by real people. Speech synthesis is discussed in more detail in Chapter 14.

Many screen readers are American in origin and the speech output from the speech synthesiser often has a pronounced American twang. Furthermore current technology in speech synthesis can sound very artificial and unnatural. This may present difficulties for non-American users and particularly for users with any degree of hearing impairment. Screen readers are discussed fully in Chapter 13.

12.6.2.4 Currently Available Screen Readers and Magnifiers

This subsection discusses some of the features of currently available products:

- The ZoomText 8.0 Magnifier/Screenreader reads out what is being typed, as well as information pointed to by the mouse, and speaks program events. The user can control the amount of information provided. It has a facility for reading documents, webpages and emails from the parent application. It also provides $2\times$ to $16\times$ magnification.
- Keytools Screenreader 4 is a utility tool bar that allows a PC to read text in any Windows-based application.

- JAWS is a widely used screen reading package with a wide range of features. It has a multilingual software speech synthesiser. It provides output to most refreshable Braille displays in computer or Grade 2 Braille. It supports all standard Windows applications. There is a scripting language for customization with non-standard Windows applications and proprietary software, as well as new tools for easier customization without the need to write scripts.
- Kurzweil 1000 is a scanning and reading software for people who are blind or severely visually impaired. It accesses text in printed or electronic form and presents it to the user audibly. It has chiming multilevel bookmarks, an online dictionary, editing tools, skimming and summarizing.
- Kurzweil 3000 provides a reading, writing and test taking solution for people with dyslexia and other learning disabilities. Documents on the computer screen look exactly like the original printed copy. The material can be read aloud and highlighted using dual highlighting. There are a range of tools, including dictionaries, study skill tools, writing tools and test taking tools. Kurzweil 3000 is available for both Windows and Macintosh.
- Read & Write 6: textHELP! is for users with dyslexia. It has speech feedback, word prediction and phonetic spell checker, dictionary and talking calculator.

12.6.2.5 Braille Display Technology

Braille output from a computer works in a very similar way to computer speech synthesiser output. The screen reader outputs what is on the screen to a Braille display. This is a slow process and each page can take some time to be read. A positive benefit is that Braille output is usually very accurate. For example, spelling mistakes in the screen information are more likely to be identified in Braille than through a slight mispronunciation from a speech synthesiser.

A Braille display sits in front of the keyboard or notetaker. This allows the user to move quickly between inputting text from a QWERTY keyboard and reading the Braille display. Full sized Braille lines are 80 characters long; this is the number of characters that appear on a computer screen. Shorter Braille lines are available and these are usually used with Braille notetakers. Most Braille displays have a router button. This allows the user to determine and change the position of the cursor on the screen. Some displays give additional detailed information about the cursor, indicating its current location by row and column. Braille display units often fit under the standard keyboard, to create a Braille Terminal. It should be noted that a refreshable Braille display is an output device, whereas a Braille keyboard is an input devices. A modern refreshable Braille display unit is shown in Figure 12.20.

The refreshable display usually has 20, 40 or 80 Braille cells each with 6 or 8 nylon or metal pins. As the characters on the computer screen are read these pins are moved up and down to represent the character read. On the screen 80 characters of text in one line may be presented to the Braille reader using an 80-cell Braille refreshable unit. Such a display can be used by Braille readers to progress word processing, database and spreadsheet applications.

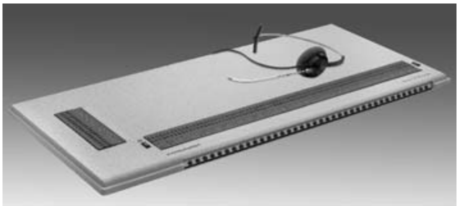


Figure 12.20. Refreshable Braille display unit (Photograph of the BRAILLEX® EL 2D-80 Braille Display unit reproduced by kind permission of Papenmeier Germany)

There are several types of Braille displays available ranging from electronic Braille displays to Braille notetakers. Braille notetakers are usually light and portable. They usually output text as speech or Braille to allow the user to check the notes taken. They either have a QWERTY or Braille keyboard. Electronic Braille displays are usually larger than notetakers. They are often equipped with screen readers that are compatible with programs such as Windows. Some of these displays are designed for use with laptop computers.

Once text has been input in Braille and saved, then it can be transferred to an ordinary PC and printed out as text using Braille translation software. This software includes the rules for producing Braille contractions and the formatting protocols required to print the Braille document. It is often necessary to pre-process the text document before the Braille transcription. Parts of the text that may require modification include tables, diagrams and pictures; the latter may require replacement by text descriptions to make them accessible. The layout of Braille documents has rules and conventions that cover positioning, the hierarchy of sections, use of running headings and page numbering. The text-to-Braille transcription software is designed to automate all these conventions as well as perform the actual text-to-Braille transcriptions.

The translation software can generally also queue the Braille document to a printer designed to emboss Braille character cells. Typical Braille embossing printers are shown in Figure 12.21. Embossing technology has print speeds vary-

Figure 12.21. Braille embossing printer (photograph of the Braille Blazer by kind courtesy of Freedom Scientific, USA)

ing from 50 to 200 characters per second. Some notetakers can be connected to a printer and print text output obtained directly.

Braille translation software includes the Duxbury Braille translater for MS Windows which can translate from text to Braille or from Braille to text. Creating a Braille document from text is its most common use. It has drivers for all the commercially available Braille embossers and can print black and white 'simulated Braille' graphics on most printers supported by Windows. It can format documents and be used as a Braille editor. It can be accessed with a mouse, speech, Braille or keyboard and be used by both blind and sighted people. Text-to-Braille software is also used to produce a wide range of documents, including official letters, bills and timetables, since there are few skilled Braille transcribers available.

12.6.2.6 Tactile Touchscreens

Touchscreens and touchpads are tactile responsive displays that allow the computer user to control processes and operations by touching the monitor screen. They allow the user to navigate the icons of typical computer menus and graphical user interfaces by touching the screen rather than using a keyboard. They are becoming increasingly popular in applications such as purchasing rail tickets and obtaining information. Although the data is input by touch, vision is required to locate the icons. Public information and ticket purchase systems can be made accessible to some blind and visually impaired people by good design with a well organised layout with good contrasts, clear, large size lettering, well spaced icons and a logical menu structure. Links should be conveniently positioned, for instance to one side of the screen and labelled with their function rather than 'click here'. Public touchscreen systems can be made accessible to additional blind and visually impaired people by the addition of a screen reader. However, visually impaired people who do not have good motor control may encounter difficulties in using the system.

12.6.3 Computer-based Reading Systems

An important application of accessible computer systems is as a reading system for text documents for visually impaired, blind and deafblind people (Hersh and Johnson 2005). This is achieved by the selection of accessible input and output devices appropriate for the individual enduser from the range of possibilities shown in Figure 12.18. This computer-based approach is in contrast to the stand-alone reading systems described in Section 12.4.1. A typical architecture and global selection of components is shown in Figure 12.22 and individual users are then able to select an appropriate subset of options from this global framework of options.

In Figure 12.24, a dotted outlined box contains those components that are housed by the computer system. A comparison of Figures 12.10 and 12.22 shows that the main principles and some of the components in the stand-alone and computer-based reading systems are very similar. The difference is that the stand-alone reading system is carefully engineered as a compact integral unit with the reading system as its main application, whereas an accessible computer system is a multi-function device with its use as a reading system being just one of many applications.

This multifunctionality is probably the main advantage of the computer-based reading system. However, in some cases, expert input may be required to configure the accessible computer system for particular users. An important application is Internet access, which is discussed in Section 12.7.

12.6.4 Accessible Portable Computers

The discussion of the accessible computer used a generalised architecture of input and output devices to make computing systems accessible to visually impaired and blind end-users. The architecture is shown in Figure 12.18 and can be seen to comprise visual, audio and tactile access methods for both the inputs and outputs to the system. Thus, it is possible by careful configuration and selection to make computer systems accessible to the wide range of end-users in the visually impaired and blind community.

Portable computers, such as the ubiquitous laptop personal computer, can also be made accessible using the methods described earlier and captured in the general architecture of Figure 12.18. However, just as laptop computers are carefully engineered and ergonomically designed for portability, computers that use tactile input and output devices have also received careful engineering design input to enhance portability and improve on the Braille notetaking type of device.

The architecture of Figure 12.18 shows that tactile access for a personal computer uses a Braille keyboard for input, and a refreshable Braille display unit for output. The PAC Mate range of portable personal computers from Freedom Scientific illustrate the engineering of tactile access, as well as a number of other options.

The range, shown in Figure 12.23, has two basic units, which use a QWERTY keyboard (top left hand unit in Figure 12.23; the QX 400 model series) and a Braille keyboard (top right hand unit in Figure 12.23; the BX 400 model series) for input. The QWERTY keyboard has 88 alphanumeric and function keys, whilst the Braille keyboard has 8-key Braille, 8 function keys, and a cursor cross key.

Both of these basic models also have voice recorder input and output *via* an audio speaker and/or a headphone jack socket. Tactile output is then achieved by the addition of a modular Braille unit, either a 20 cell or a 40 cell refreshable Braille

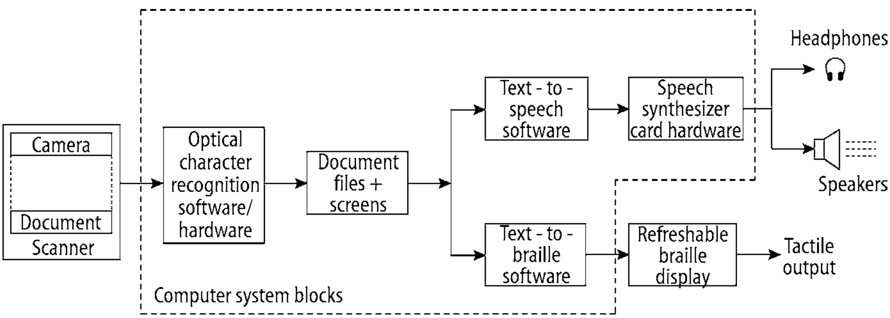


Figure 12.22. Reading system architecture based on the accessible computer

Figure 12.23. Range of accessible portable computers (photograph of the PAC Mate™ computer range by kind courtesy of Freedom Scientific, USA)

display unit. The use of the Braille display can be seen in the remaining four units of Figure 12.23. These machines support a wide range of software and use the JAWS® screen reader software.

12.7 Accessible Internet

Being able to access and use the Internet is becoming increasingly important for full participation in society. There are moves toward the information society, with all facilities and services available on-line. The term 'digital divide' is often used to describe the increasing gap between those who do and do not have access to information and communications technology. As currently envisaged, the information society has a number of disadvantages and may lead to increasing inequalities. However, discussion of the social role and impacts of information technology is outside the scope of this book. What is relevant here is that disabled people are able to have the same quality and extent of access to and use of information technology as non-disabled people. Currently this is not the case.

In particular, many websites have been designed in ways that are difficult to navigate or are inaccessible to tools like screen readers. Sites that offer services like shopping or banking need to be carefully designed for access by visually impaired and blind people. Since these sites are popular and often provide an essential service to those who experience serious barriers to leaving the house, inaccessibility can be very frustrating and discriminatory for potential visually impaired and blind customers. There are also disadvantages for the firms and other organisations that do not make their websites accessible including loss of custom and the perception of poor public relations and service levels. Thus, firms and organisations could benefit from making their websites accessible.

12.7.1 World Wide Web Guidelines

There are a number of different initiatives to make the Web more accessible. One of the best known is the World Wide Web Consortium (W3C Web Accessibiliy Initiative), which was established in October 1994 to help make the Web more accessible to everyone. The W3C is currently trying to do this by developing common protocols that promote the evolution and ensure the interoperability of the Web.

The W3C is involved in four domains: architecture, user interface, technology and society, and the Web Accessibility Initiative (WAI). The W3C set up the Web Accessibility Initiative (WAI) to ensure that the Web is accessible to everyone, including disabled people and people using older computers and older version web browsers. It is working in five key areas: technology, guidelines, tools, education and outreach, and research and development.

The Web Content Accessibility Guidelines aim to enable the creation of webpages that are accessible to everyone, including disabled people, people using older computers and older version browsers and people working in noisy or otherwise unfavourable environments. In some countries, webpage accessibility is required by legislation on accessibility to disabled people. In addition, webpages are generally intended to have as wide a distribution as possible, so there are clearly benefits in including disabled people and people with older computers and browsers and disadvantages in excluding them.

People who have a sensory impairment often use assistive technologies when using the Internet. These technologies are very helpful, but require appropriate webpage design to allow access. When the assistive technology finds something on the page it cannot deal with, the page becomes inaccessible to the user. Examples of such problems include unlabelled graphics, badly positioned frames and tables within a page.

The guidelines are intended to be simple and straightforward to use and the grouping into four categories in the second version may make them easier to remember. They suggest simple common sense standards such as providing a textual description of a graphic or a sound through the "alt-text" or "long-desc" tags. This should also benefit people who have a slow Internet connection or who prefer not to download graphics, as well as blind people. If people are working in a noisy environment they may not be able to hear sound clips so providing a text alternative would be of benefit to them, as well as to deaf people and people who prefer text or find it easier to understand than graphics.

The use of guidelines, such as those drawn up by the W3C, has a number of advantages. In particular it provides a structured and easily followed approach to designing webpages to make them accessible. This makes it much easier to identify potential and actual problems and develop technologies to resolve them.

However, some web designers may feel restricted by the guidelines in what they can put on the webpage and therefore discouraged from using them. For example, designers may feel that they cannot use graphics on webpages in case a visually impaired person tries to access their page. This is not the case. The aim is not to restrict the creativity of web designers, but to ensure that they extend this creativity to designing webpages which can be accessed by everyone (while remaining interesting and attractive). Thus, the designer simply needs to consider what graphics are to be put on the page and provide a description to accompany the graphics. Particular care needs to be exercised when using tables and frames. It should be noted that older screen readers may have problems reading them.

12.7.1.1 W3C Accessibility Guidelines 1.0

WCAG 1.0 was the first version of the guidelines, which was published in 1999. It comprises 14 guidelines based on the following two themes:

1. *Ensuring graceful transformation* This means that the pages remain accessible for people with physical, sensory and cognitive impairments and despite any limitations of the environment they are working in, such as noise, and technological barriers, such as using old computers and old version browsers. Guidelines 1–11 address this theme. Key principles include the following:
   - Separating structure from presentation.
   - Providing text, including text equivalents, as text can be made available to almost all browsing devices and accessible to almost all users.
   - Creating documents that can be used by blind and/or deaf users by providing information to other sensory modalities. This does not mean a prerecorded audio version of a whole site, as blind users (who are not also deaf) can use screen readers to access text information.
   - Creating documents that do not rely on a particular type of hardware, so they can be used by people who do not use a mouse, with small, low resolution, black and white screens, who use voice or text output.
2. *Making content understandable and navigable* Language should be clear and simple and the mechanisms for navigating between and within pages should be comprehensible. Navigation tools and orientation mechanisms should be provided in pages to maximise usability and accessibility. This is particularly relevant to users accessing pages one word at a time through speech synthesis or Braille display or one section at a time through a small or magnified display. Guidelines 12–14 address this theme.

Each of the guidelines has a number of associated checkpoints. The checkpoints have been assigned three priority levels by the Working Group based on their impact on accessibility:

Priority 1. Content must satisfy this checkpoint or one or more groups will find it impossible to access the information. Satisfying this checkpoint is a basic requirement for some groups to use web documents.

Priority 2. Content should satisfy this checkpoint or one or more groups will find it difficult to access information in the document. Satisfying this checkpoint will remove significant barriers to accessing web documents.

Priority 3. Content may address this checkpoint or one or more groups may find it somewhat difficult to access information in the document. Satisfying this checkpoint will improve access to web documents.

A number of suggested techniques for achieving it are stated for each checkpoint. There are three levels of conformance: A, Double-A and Triple-A which comprise satisfaction of all Priority 1, all Priority 1 and 2, and all Priority 1, 2 and 3 checkpoints respectively. Conformance claims should specify the guidelines title, its URI, the conformance level, *e.g.* A and the page site or defined portions of a site covered by the claim. Alternatively one of the three W3C icons can be used on each conforming page with a link to the appropriate W3C explanation.

The fourteen guidelines are as follows:

Guideline 1. Provide equivalent alternatives to auditory and visual content.
Guideline 2. Do not rely on colour alone.
Guideline 3. Use mark-up and style sheets and do so properly.
Guideline 4. Clarify natural language usage.
Guideline 5. Create tables that transform gracefully.
Guideline 6. Ensure that pages featuring new technologies transform gracefully.
Guideline 7. Ensure user control of time-sensitive content changes.
Guideline 8. Ensure direct accessibility of embedded user interfaces.
Guideline 9. Design for design independence.
Guideline 10. Use interim solutions.
Guideline 11. Use W3C technologies and guidelines.
Guideline 12. Provide context and orientation information.
Guideline 13. Provide clear navigation mechanisms.
Guideline 14. Ensure that documents are clear and simple.

For illustration, the checkpoints and techniques for the first two guidelines will be stated. Guideline 1 has the following five checkpoints, with checkpoints 1.1–1.4 Priority 1 and 1.5 Priority 3:

Checkpoint 1.1. Provide a text equivalent for every non-text element, for instance using ‘alt’ and ‘longdesc’ in element content. Non-text elements includes images, graphics, symbols, image map regions, animations, applets, program objects, frames, scripts, audio files and video.

Checkpoint 1.2. Provide redundant text links for each active region of a service-side image map.

Checkpoint 1.3. Provide an auditory description of the important information of the visual track of a multimedia presentation until user agents can automatically read out its text equivalent.

Checkpoint 1.4. Synchronise equivalent alternatives, such as captions or auditory descriptions of the visual track with the presentation for any time-based multimedia presentation, such as a video or animation.

Checkpoint 1.5. Provide redundant text links for each active region of a client-side image map until user agents are able to provide text equivalents for client-side image map links.

Core techniques for achieving Checkpoint 1.1 include the use of text equivalents, whereas HTML techniques include the following:

- Images used as bullets and graphical buttons.
- Text for images used as links.
- Short text equivalents, such as 'alt-text', and long descriptions of images.
- Text (and non-text) equivalents for applets, programmatic objects, multimedia and client-side image maps.
- Describing frame relationships and writing for browsers that do not support FRAMEs.
- Alternative presentation of scripts.

12.7.1.2 W3C Accessibility Guidelines 2.0

A revised form of the Web Content Accessibility Guidelines is currently being developed in consultation with users and drawing on the experience with WCAG1. The Web Content Accessibility Guidelines 2.0 (WCAG 2) are still in draft form and may be updated or replaced by subsequent documents. The final agreed version will be published as a W3C Recommendation, but until then WCAG 1.0 is the stable reference version. Both WCAG 1.0 and 2.0 have the same overall goal of promoting accessibility of web content, and WCAG 2.0 has the following additional aims:

1. Ensuring that the requirements can be applied across technologies.
2. Ensuring that the conformance requirements are clear.
3. Ensuring that the deliverables are easy to use.
4. Addressing a more diverse audience.
5. Clearly identifying who benefits from accessible content.
6. Ensuring that WCAG 2.0 is compatible with WCAG 1.0.

WCAG 2.0 is intended to be more efficiently organised than WCAG 1.0 and to incorporate the experience and errata from WCAG 1.0. In addition, it may adjust the priority of some checkpoints and modify, remove or add requirements to take account of changes in web technologies since the publication of WCAG 1.0. The majority of WCAG 1.0 checkpoints can be mapped into WCAG 2.0 success criteria (http://www.w3.org/WAI/GL/2005/11/23-mapping.html). However, WCAG 2.0 does incorporate a number of additional guidelines, such as the provision of sign language interpretation for multimedia.

The WCAG 2 guidelines are organised into the following four accessibility principles:

Principle 1. Content must be perceivable.

Principle 2. User interface components in the content must be operable by each user.
Principle 3. Content and control must be understandable to each user.
Principle 4. Content must be robust enough to work with current and future technologies.

Each of these principles has a list of guidelines that address the principle and each principle has a number of success criteria. The success criteria are grouped into three levels of conformance, representing increasing levels of accessibility for the associated guidelines. They are statements that will be true or false according to whether or not the specific web content meets the criteria. The principles, guidelines and success criteria are intended to be applicable to all web-based content and to be independent of the technology used. There are three levels of conformance with the success criteria. Level 1 and level 2 success criteria respectively give a minimum and an enhanced level of accessibility through markup, scripting or other technologies that interact with or enable access, including assistive technologies. Both levels 1 and 2 are intended to be applied to all web resources. Level 3 success criteria enhance accessibility for disabled people and are not applicable to all web resources.

Some guidelines do not have level 1 or level 2 success criteria. The approach to conformance has changed from the previous version WCAG 1.0, in which checkpoints were assigned Priorities 1–3 according to their impact on accessibility, with Priority 3 checkpoints less important than Priority 1 ones. In WCAG 2.0 this is no longer the case and all the criteria are essential for some groups of users.

All the success criteria are intended to be testable by either computer programs or people who understand WCAG 2.0. Tests of this type can be used to determine whether or not the content conforms to the criteria. However, there is some anecdotal evidence that meeting the WCAG 1.0 guidelines has not always been sufficient to resolve all end-user problems and the same may be the case with WCAG 2.0. Therefore, there are advantages in testing being carried out by disabled end-users, though this may not always be feasible.

The guidelines and criteria do not require or prohibit the use of any specific technology, as long as it is supported by accessible *user agents*, which are defined as software which retrieves and makes available web content for users. This includes web browsers, media players, plug-ins and assistive technology, including screen readers, screen magnifiers and alternative keyboards. However, a particular *baseline* or set of technologies supported by and active in accessible user agents may be specified for a given web content. In this case the criteria only need to be checked for technologies in the baseline. Web developers can use non-baseline technologies under the following conditions:

- All content and functionality can be accessed using only technologies in the specified baseline.
- The non-baseline technologies do not impede access to the content when used with user agents that either support only baseline technologies or support both baseline and non-baseline technologies.

There are three levels of conformance when the user agents support only the technologies in the chosen baseline:

1. Level A: all level 1 success criteria are met.
2. Level AA: all level 1 and level 2 success criteria are met.
3. Level AAA: all levels 1, 2 and 3 success criteria are met. However, the working group is considering modifying AAA conformance, as not all level 3 web criteria can be applied to all web content and some are not necessary in some circumstances.

The term 'delivery unit' rather than 'page' is used in WCAG 2.0, as it also covers web applications. Claims of conformance are stated in terms of delivery units and must include the following information:

1. The date of the claim.
2. The guidelines title and version.
3. The URI of the guidelines.
4. The conformance level satisfied, *i.e.* level A, AA or AAA.
5. The baseline, for instance, in terms of the individual baseline technologies.
6. The delivery units covered.

The first principle *Content must be perceivable* has the following four guidelines:

1.1. Provide text alternatives for all non-text content.
1.2. Provide synchronised alternatives for multimedia.
1.3. Ensure that information, functionality and structure can be separated from presentation.
1.4. Make it easy to distinguish foreground information from background images or sounds.

The second principle *User interface components in the content must be operable by each user* has the following five guidelines:

2.1. Make all functionality operable *via* a keyboard interface.
2.2. Allow users to control time limits on their reading or interaction.
2.3. Allow users to avoid content that could cause seizures due to photosensitivity.
2.4. Provide mechanisms to help users find content, orient themselves within it, and navigate through it.
2.5. Help users avoid mistakes and make it easy to correct them.

The third principle *Content and controls must be understandable to each user* has the following two guidelines:

3.1. Make text content readable and understandable.
3.2. Make the placement and functionality of content predictable.

The fourth principle *Content must be robust enough to work with current and future technologies* has the following two guidelines:

4.1. Use technologies according to specification.
4.2. Ensure that user interfaces are accessible or provide accessible alternative(s).

For illustrative purposes, the success criteria for the guideline 1.2: *Provide synchronised alternatives for multimedia* will now be stated:

Level 1: 1.2.1. Captions, *i.e.* synchronised transcripts of dialogue and important sound effects are provided for pre-recorded multimedia.
1.2.2. Audio descriptions of video are provided for pre-recorded multimedia.
Level 2: 1.2.3. Real-time captions are provided for live multimedia.
Level 3: 1.2.4. Sign language interpretation is provided for multimedia.
1.2.5. Extended audio descriptions of video are provided for prerecorded multimedia.

The other guidelines have varying numbers of associated success criteria and do not always include success criteria at all three levels. The guidelines document is supported by the document Understanding WCAG 2.0 (UWCAG, undated) which also includes discussion of the benefits of the different success criteria and ways of meeting them with examples, common mistakes and how to avoid and correct them.

12.7.2 Guidelines for Web Authoring Tools

Authoring tools are often used to create web content. They determine how the content is implemented and can have a significant impact on determining whether or not web content meets the Accessibility Guidelines. Therefore authoring tools should conform to the Authoring Tool Accessibility Guidelines.

The term *authoring tool* refers to software that is used to create web content, including:

- Editing tools for web content, such as HTML and XML editors.
- Tools that can be used to save material in a web format, such as word processors.
- Tools that transform documents into web formats.
- Tools to produce multimedia to be used on the Web.
- Tools for site management or publications.
- Tools for the management of layout, such as cascading style sheets (CSS) formatting tools.

The Authoring Tool Accessibility Guidelines 1.0 (ATAG, undated) are currently being updated to produce Guidelines 2.0. The aim of the guidelines is to produce authoring interfaces which are accessible to disabled authors, and to support and promote the production of accessible web content by all authors. It is equally important that people are able to produce their own accessible webpages as well as access other people's webpages.

There are seven guidelines in ATAG1. Each guideline has a number of checkpoints with one or more associated priority levels. Some checkpoints that refer to generating, authoring or checking web content have multiple priorities, depending on the priority of the associated checkpoint in the Web Content Accessibility Guidelines 1.0. The priorities are defined in terms of the extent to which they meet the following goals for the authoring tool, with Priorities 1, 2 and 3 respectively essential, important and beneficial to meeting these goals:

- Accessibility of the authoring tool.
- Production of accessible content by the authoring tool by default.
- Encouraging the creation of accessible content.

There are again three conformance levels, A, Double-A and Triple-A, relating to satisfaction of all Priority 1, all Priority 1 and 2, and all Priority 1, 2 and 3 checkpoints, in each case including all the relative priority checkpoints. The seven guidelines are as follows:

Guideline 1. Support accessible authoring practices. Authoring tools should automatically generate accessible markup and, where appropriate, guide authors in producing accessible content. Applications that convert documents from other formats to markup or make markup changes to facilitate efficient editing should not introduce inaccessible markup or remove accessibility content.

Guideline 2. Generate standard markup. Conforming to standards encourages interoperability and accessibility. The assistive technologies used with browsers and multimedia players can generally only access web documents that use valid markup, making that an essential aspect of the accessibility of authoring tools. W3C Recommendations or published standards, if there are no W3C Recommendations, should be used.

Guideline 3. Support the creation of accessible content. Structuring information and providing equivalent information in alternative formats are the basis of accessible design. Authoring tools should therefore be designed to facilitate and automate this process, for instance by prompting authors to include equivalent alternative descriptions at appropriate times.

Guideline 4. Provide ways of checking and correcting inaccessible content. Authoring tools should be designed so that they can identify inaccessible markup, if possible automatically, and enable its correction even when the markup is hidden from the author. Authoring tools support should allow a number of different authoring styles, as authors are more likely to use accessible authoring practices if they can configure the accessibility features to support their standard work patterns. For instance, some authors may prefer to be alerted to accessibility problems as they occur, whereas others would rather make a check at the end of each editing session.

Guideline 5. Integrate accessibility solutions into the overall 'look and feel'. Creating accessible content should be a natural and easy to follow process with an authoring tool, as authors as less likely to accept features that are not properly integrated. In addition the relative prominence of different ways of accomplishing a task can affect which one the author chooses.

Guideline 6. Promote accessibility in help and documentation. Many web authors are not familiar with web accessibility issues and therefore require explanations of accessibility problems and examples.

Guideline 7. Ensure that the authoring tool is accessible to disabled authors. Authoring tools are software programs with standard user interface elements and need to take account of user interface accessibility guidelines. Any custom interface components should be accessible through the standard access mechanisms so that assistive technologies can be used with them. The tools must also enable authors to edit using their stylistic preferences and publish in different styles. The style preferences of the editing view (used by the author to edit the document) should not affect the markup of the published document. For instance, visually impaired authors may want to edit using a larger font size than the default document size. Authoring tools should also enable all authors to navigate documents efficiently when editing and, for instance, using screen readers, Braille displays, screen magnifiers or switches. Therefore authoring tools should provide an editing view that gives a sense of the overall structure and allows structured navigation.

12.7.2.1 Authoring Tool Accessibility Guidelines 2.0

In Authoring Tool Accessibility Guidelines 2.0 (ATAG 2.0) an *authoring tool* is defined as 'any software, or collection of software components, that authors use to create or modify web content for publication.' A collection of software components are software products which are used together or separately. The main aim of ATAG 2.0 is to ensure that 'Everyone should have the ability to create and access web content'. The design of web authoring tools determines who can use them to produce web content and the accessibility of the resulting web content determines who can use it. ATAG 2.0 again comprises seven guidelines, but they are now divided into two parts. Part A comprises guidelines and checkpoints for making authoring user interfaces accessible and part B the guidelines for supporting the creation of accessible web content. WCAG is used in ATAG 2.0 as the benchmark for determining the accessibility of web content and web-based authoring interfaces, as well as for defining the terms accessible web content and accessible authoring interface. As with ATAG 1.0 there are again three levels of conformance, A, Double-A and Triple-A, with slightly different definitions from those in ATAG 1.0. Levels A, Double-A and Triple-A again require satisfaction of all regular checkpoints of Priority 1, Priority 1 and 2, and Priority 1, 2 and 3 respectively, but require satisfaction of all the relative priority checkpoints to at least level 1, 2 and 3 respectively.

Part A: Make the user interface accessible comprises the following four guidelines:
Guideline A.1: Authoring interface must be perceivable.

Guideline A.2: Authoring interface must be operable.
Guideline A.3: Authoring interface must be understandable.
Guideline A.4: Authoring interface must be access system friendly.

Part B: Support the production of accessible content comprises the following three guidelines:
Guideline B.1: Enable the production of accessible content.
Guideline B.2: Support the author in the production of accessible content.
Guideline B.3: Promote and integrate accessibility solutions.

For illustration the checkpoints associated with the success criteria for Guideline A.1 will be stated:

A.1.1. Provide text alternatives for all non-text content in the user interface (Priority 1). This will give access to people who have difficulty perceiving non-text content through a text alternative, such as Braille. The success criteria are as follows:

1. All user interface non-text objects used to convey information, such as a toolbar icon or sound effect, must have a text alternative, for instance a text label or long text description.
2. All editing views must always include an option to display the available text alternatives for non-text objects in the content.

A.1.2. Provide synchronised alternatives for multimedia in the user interface (Priority 1). This allows people who have difficulty accessing multimedia information to have it made available by other means, for instance audio descriptions of visual information for blind and visually impaired people. The success criteria are as follows:

1. All user interface multimedia used to convey information, such as tutorial videos, must have synchronised alternatives, for instance captions and audio descriptions.
2. All editing views must always include an option to display the available text alternatives for multimedia in the content.

A.1.3. Ensure that all displays are configurable (Priority 1). This allows alternative display configurations to use the authoring interface. The success criteria are dependent on the capabilities of the platform, which includes operating systems and browsers. The success criteria are as follows:

1. The authoring tool must provide at least the same configurable properties with the same configuration ranges as the platform when the visual display (fonts, sizes, colours, spacing, positioning) is controlled by the authoring tool rather than the platform.
2. The authoring tool must provide the same configurable properties with at least the same configuration ranges as the platform when the audio display (volume, voices) is controlled by the authoring tool rather than the platform.

A.1.4. Allow the display preferences for the editing view to be changed without affecting the document markup (Priority 1). This allows authors to use different preferences to view and control the document from those in the published version. The success criteria are as follows:

1. The author must be able to configure the presentation settings of editing views without affecting the web content being edited.

A.1.5. Ensure that information, functionality and structure can be separated from presentation (Priority 1). Separating content and structure from presentation allows different authors to use different interfaces and authoring tools without losing any information or structure. The success criteria (which are being revised) are currently as follows:

1. Information conveyed by variation in the presentation of text, *e.g.* by spatial location, must also be conveyed in text or made available programmatically.
2. Information conveyed by colour must also be conveyed in text or made available programmatically and conveyed in a way that is visually evident when colour is not available, for instance by shape.
3. If content is structured the structure must be made available programmatically.
4. If the sequence of content affects its meaning, the sequencing information must be made available programmatically.

12.7.3 Accessible Adobe Portable Document Format (PDF) Documents

Adobe PDF documents are frequently used on the Web and to distribute electronic documents over corporate networks, by email and digital media. Therefore, it is important that PDF documents are fully accessible. Adobe has produced a document on using the most recent version of its software, Adobe® Acrobat® 7.0, to create accessible PDF documents (ADOBE, 2005a), as well as information on the accessibility features of versions 6.0 and 7.0. This includes a list of features for accessible PDF documents, which could be taken as accessibility guidelines. However, there does not seem to be any information on PDF accessibility produced independently of the manufacturer or any surveys of blind and visually impaired people to investigate their experiences with PDF documents other than the small-scale survey performed by the authors (Hersh and Johnson 2006). In addition, a number of accessibility features are only available in the Professional version, whereas many users only have Adobe® Reader® or Acrobat Standard. This then puts a practical limitation on the likelihood of PDF documents being fully accessible.

Accessible PDF documents have the following characteristics (ADOBE, 2005b):

1. The document is a searchable text file rather than an image-only scan. Therefore scanned document files require the application of optical character recognition to make the image produced by scanning into searchable text with selectable graphics. Both Acrobat 7.0 Professional and Standard are able to convert scanned images to searchable PDF documents.
2. Any form fields should be accessible. If the PDF document is a form to be completed, then the interactive (fillable) form fields and their descriptions

need to be readable by screen readers and there should be a preset tab order to facilitate navigation among the form fields. Acrobat 7.0 Professional and LiveCycle® Designer have these features.

3. The document structure is indicated by tags. Accessible Adobe PDF documents use *tags* to indicate structural elements of a document, such as titles, headings, figures, text and tables, and the relationship between these elements. Although different assistive technologies may process document structures differently, using a consistent tagging system generally improves accessibility for disabled people. Some documents can be prepared for tagging before conversion to PDF or tags can be added to documents in PDF form.
4. The reading order should be clear and easy to follow. Assistive technology is designed to read page content in the order the content is received from an application. Therefore if the order is not logical, for instance if a heading is after its text or a figure caption is separated from the figure description, disabled people may not be able to make sense of the content. A structured reading order is generally produced by tagging. However, the reading order may need to be corrected in complex documents and this generally requires Acrobat 7.0 Professional.
5. A descriptive text is provided for all graphics, links and form fields. Accessible Adobe PDF documents have descriptive or alternate text to describe illustrations, graphs, charts form fields and links, as screen readers are not able to read graphical elements. The alternate text enables screen readers to interpret them and read a description to the user. Descriptive text about screenreaders and URLs, which can be read by screen readers, facilitates navigation. Acrobat 7.0 Professional is required to add alternate text and descriptions to PDF pages.
6. Navigational aids should be provided. Accessible PDF documents provide navigational aids in the form of links, bookmarks, useful and frequent headings, a detailed table of contents and an optimised, preset tab order for forms and embedded links. Users can use these navigational aids to go directly to a particular place in the document without having to read the whole document. Most navigational aids can be set during conversion from authoring applications to PDF. Bookmarks and links can be set using Acrobat 7.0 Professional or Standard.
7. The document language is specified to enable multilingual screenreaders to switch between languages during operation. Both Acrobat 7.0 Professional and Standard can set the document language for the whole document, whereas the Professional version is required to set different languages for different parts of a multi-language PDF document.
8. The fonts used allow characters to be extracted to text. This requires fonts to contain sufficient information to enable Adobe Reader or Acrobat to extract correctly all the characters to text. For instance, characters are extracted to text when users listen to the text using a screen reader or the Read Out Loud tool in Adobe Reader or Acrobat or when they copy, paste or save text to a file. If there is insufficient font information, incorrect output may be received, with

words or characters omitted or question or other marks added when a PDF file is copied, pasted or saved as text. Some fonts are inaccessible in the sense that they do not contain sufficient information for Adobe Reader or Acrobat to correctly extract all the text characters. These documents can be read on the screen, but cannot be read by screen readers and therefore such fonts should be avoided. Only Acrobat 7.0 Professional is able to check for the presence of inaccessible fonts.

9. The security settings should not interfere with screen readers. Settings to restrict copying of part of a PDF documents could interfere with screen readers, since they carry out a type of copying in order to read the text back. Acrobat 7.0 Professional and Standard both have an option which resolves this problem, called 'Enable text access for screen reader devices for the visually impaired'.

Acrobat 6.0 and 7.0 and the free Adobe Reader also have a number of accessibility features including the following (ADOBE 2005a,b):

- Synthesis of the text into speech so it can be read aloud to enable the reading order to be checked. Although a screenreader is not required, it is advisable particularly for complicated texts, as it provides better navigation, allows users to toggle between table and text reading modes and allows access to PDF forms as well as documents. Acrobat 6.0 uses Microsoft Active Accessibility, an application programming interface which provides information about the content and user interface of Windows-based programs to assistive technologies using speech and/or refreshable Braille displays. Since Microsoft Active Accessibility is widely supported, Acrobat 6.0 can be used with a number of different screen readers and other assistive technologies.
- Evaluation of the accessibility of a PDF file so that blind and visually impaired people know whether it is accessible and can be read in the correct order. This is done using the Reader Quick Check Feature which can indicate whether the document is a scanned image or tags are present.
- Customising the font size in the navigation panels. Increasing the font size of the navigation panels can enable visually impaired people to read bookmarks, comments and signatures in a PDF file.
- Choosing alternate reading orders. Improving the reading order of text and form fields can be important to users of screen readers in documents that have not been optimised for accessibility. There are three options:
  - The default is allowing Acrobat to determine the reading order using columns, boundaries, form fields and other layout information in the PDF file.
  - Acrobat uses the word order in the PDF document's print instructions.
  - Acrobat reads the page from left to right and top to bottom.
- The ability to view documents in high contrast mode to improve readability. Colour contrast can be increased in Acrobat 6.0 by replacing the specified colours in the document with custom colour schemes created by the user or high contrast colour settings defined in the operating system.

- Zooming in on the text and reflowing it to fit any size view. Viewing tagged PDF files using the Acrobat 6.0 reflow feature allows users to use large type display and the text to automatically reflow to fit the available screen space. This is preferable to magnification and manual scrolling.
- Saving PDF content as text to use with a screen reader that is not compatible with Microsoft Active Accessibility, enabling the opening of files containing tables or other content that is too complicated for the Read Aloud function and sending text to a Braille printer.
- Tagging of existing PDF documents using Acrobat 6.0 to enable better performance with a screenreader.

Some features require the use of Acrobat Professional 6.0 or 7.0. These include identifying accessibility problems and automatically inserting each form field into the tag to allow the recognition of form fields, identification of proper reading order and reading any descriptive text.

12.7.4 Bobby Approval

The Center for Applied Special Technology (CAST) was set up in 1984 to develop innovative technology-based educational resources and strategies based on 'Universal Design for Learning', which requires the following:

- *Multiple means of representation*, to give learners various ways of acquiring information and knowledge.
- *Multiple means of expression*, to provide learners with alternative ways of demonstrating what they know.
- *Multiple means of engagement*, to take account of learners' interests, offer appropriate challenges, and increase motivation.

In 1996, CAST created a webpage evaluation tool called Bobby to encourage webpage designers to adhere to the W3C accessibility guidelines. The concept is that the Web should be a resource for all to use. A free version, called WebXACT, is available (WebXACT, undated), but can only be used to test single pages of web content for quality, accessibility and privacy issues by entering the URL of the page. Testing a website consisting of several pages would require the URL of each page to be entered separately. The full version was commercialised when Watchfire acquired Bobby from CAST in 2002 and took over the responsibility for development, marketing and distribution of the technology. The current version can be purchased as Watchfire Bobby 5.0 (WB, undated). It is intended to be a comprehensive web accessibility tool and encourage compliance with both the W3C Web Content Accessibility Guidelines and Section 508 of the US Rehabilitation Act.

Bobby spiders through a website and tests each page for compliance with accessibility requirements, including readability by screen readers, the provision of text equivalents for all images, animated elements, audio and video displays. Bobby can see local webpages, as well as webpages behind firewalls. It performs over 90 accessibility checks. During a scan, Bobby checks HTML file against selected

Figure 12.24. Big Button 100 telephone (photograph reproduced by kind permission of British Telecommunications plc, UK)

accessibility guidelines and then reports on the accessibility of each webpage. It can integrate with HTML editors to facilitate the fixing of identified problems. Sites that pass the test can display the Bobby Approval logo on the site, the head and helmet of a U.K. police officer or 'bobby'.

12.8 Telecommunications

Although the functions of speaking and listening on the telephone do not require vision, dialling a number does. The additional telephone functions that are now available, such as caller identification and sending text messages, also require vision. Therefore, these functions need to be made accessible to blind and visually impaired people. For dialling, the simplest solution is to have large keys with very clear contrast characters and a small bump to mark the number 5, as shown in Figure 12.24, though this is more practicable for landlines than the small handsets of mobile phones. There are also voice dialling and talking caller identification systems. Screen readers for mobile and landline phones have been developed.

12.8.1 Voice Dialling General Principles

Voice dialling allows the user to dial a number by speaking it or the name of the person being called. There are two types, speaker dependent and speaker independent. In speaker dependent systems the telephone will only respond to recorded names, and generally only when spoken by the person who recorded them initially. Therefore, each name must be spoken and recorded one to three times first before it can be dialled. The maximum number of voice dial entries is generally limited. Speaker independent voice recognition does not require pre-recording and the name can be spoken by anyone. The system will automatically match the spoken name with the closest name in the telephone book. Speaker dependent voice dialling is more common, particularly on older telephones.

Speaker independent voice recognition systems are required to connect a single word to a number in the phone book. This involves sampling an analog speech signal and detecting the start and end-points of the word. The three most commonly used algorithms are threshold detection, zero-crossing (that is, the number of times the signal amplitude crosses the axis) and zero-crossing with a threshold, which is preferred in noisy environments. Zero-crossing rates are generally higher in speech sequences than in noise and speech generally has higher magnitudes. The noisy speech signal is filtered through a FIR filter to improve the signal-to-noise ratio. The recorded sequence starts before the start of the word to give information on the noise that can be used in the construction of the filter. The speech signal is blocked into frames and the coefficients are extracted from each frame and quantized. Either matched filters or the Hidden Markov Model can be used to detect what labels correspond to the spoken word.

12.8.1.1 Voice Dialling Systems

There are a number of different voice dialling systems, some of which have additional features, such as screenreading, and are generally correspondingly expensive. Some dialling systems can be used with any telephone, whereas others are designed for a specific series of telephones. OnStar Personal Calling provides voice dialling features. The user is able to place calls or store telephone numbers by saying 'dial' or 'store' respectively and then saying the whole number without pauses. The system then repeats the number to check whether it is correct.

The Dialtalk Pocket Telephone Dialler (see Figure 12.25) from Cobolt Systems speaks each digit as it is entered and allows the user to erase and correct each digit. Pressing the 'speak' button reads back the whole number. The tones can then be transmitted to the telephone to dial the number by holding the Dialtalk close to the mouthpiece and pressing 'dial'. The Dialtalk has ten memories, adjustable volume and last number redial. It can be used with tone or pulse telephones that are connected to a tone exchange. Each memory can hold 15 digits and memories can be combined for dialling very long numbers. The Handsfree Freedom Voice Dialer dials numbers that have been stored by the user speaking the person's name. It can store a maximum of 50 names. The Infinity Voice Dialer will also dial a number spoken by the user that has not been programmed into the dialler. It has a maximum of 40 spoken names or numbers. In both cases the programmed phone numbers can be up to 35 digits long. Voice prompts are given for recording and deleting names and numbers. The system is speaker dependent and trained to hear a particular voice, giving better speech recognition. It works with all standard telephones; include cordless ones, as long as they are touch tone. The Voice dialler is installed between the telephone and wall jack, similarly to an ansaphone. The system may not work on business telephone systems which have their own internal wiring systems.

The Voice Activated Speaker Phone has no handset. When the telephone rings and the user says 'hello', the telephone is answered and the conversation takes place over a speaker phone. The call can be ended by pressing the 'off' button on the keypad or using the wireless remote. Compusult's TeleTalk (see Figure 12.26)

Figure 12.25. Dialtalk pocket telephone dialler (photograph reproduced by kind courtesy of Cobolt Systems Ltd., UK)

Figure 12.26. Teletalk (photograph reproduced by kind courtesy of Compusult Ltd., Canada)

provides access to the telephone console indicator and message display information for blind and visually impaired people. The system captures call related data and converts it to both speech and output on a large high contrast display. There are different TeleTalk systems for different models of telephone. There are several different configurations, based on different combinations of electronics hardware, software and computers.

Caredec Products distributes a number of talking telephone products. TALKS 80 is a screen reader for the Nokia Communicators 9x series of phones. It enables the user to write and read text messages, e-mail and text documents, dial a number from the telephone book, add entries to and edit the contact data base. It also provides an appointment calendar, a calculator and an alarm clock. Talks 60 provides similar functions for the Symbian Series 60 phones. Both products use ETI Eloquence text-to-speech software.

12.8.2 Talking Caller ID

Talking caller identification systems give blind and visually impaired people the number and, in some systems, also the name of the caller, with a database to store names and numbers. Some systems have additional features such as being able to send customised messages to particular callers, notifying the user by email that they have received a call and refusing calls from particular numbers. Most of the systems require the user to subscribe to a caller ID service.

Cobolt Systems's talking caller ID allows the user to record a message, typically the caller's name so that future incoming calls from the same person will lead to the message being spoken. The message features can be used with up to 20 different numbers. The speaker volume is adjustable. Users need to subscribe to the BT/Cable caller display service. An incoming call log displays up to 99 calls, with the time, date, number and name (if stored in the directory), but only announces the calls that have been assigned a name or other message. Assistance from a sighted person may be required in setting up the caller ID. The fact that only some calls are announced and that the device cannot necessarily be set up independently are minor disadvantages.

Talking Caller ID is a US shareware program that informs the user who is calling before they pick up the telephone. It requires users to subscribe to a telephone company caller id service and have a caller id capable modem or caller id computer hardware device. Talking Caller ID can work with both name and number or number alone services. It stores details of every person who has called, with their name, telephone number and the date, time and number of rings, as well as any custom greeting or photo, in a database. The program can look up the names in the database, allowing users with number only caller id services to see and hear the name of the caller. A recorded custom greeting in the user's voice, photo or text file can be spoken or displayed whenever a call from a particular caller is received. Installing a free speech recognition engine enables the program to support speech recognition, including training for higher accuracy for the user's voice.

Call blocking by number or name can be used to screen out unwanted calls, with a database to which details of unwanted callers can be added. When calls are received from unwanted callers, the modem picks up the line and hangs up and the program informs the user that the call was blocked and states who was trying to call. The e-mail notification feature can send an e-mail to any Internet mail address when a call is received. The program also has an ansaphone feature with customisable outgoing messages. The Remote Control for Talking Caller ID is an add-on that allows users to receive caller id information from their computer from any touch tone phone. The computer reads back the information requested through a speech engine. The program can be used in the UK, but requires appropriate hardware. The option Caller ID Plug provides an audio alert whenever the user receives new voice mail. There is also the option to send a notifying message to the user's pager or e-mail when voice mail is received.

Panasonic produce several cordless phones with Talking Caller ID and which incorporate text-to-speech programs. A Caller ID service is required.

12.8.3 Mobile Telephones

There are currently no large button style mobile telephones. However, some have two or three large function buttons and some have clearer displays than others. For instance, some visually impaired people like the font size adjustment facility on the Sony Ericsson T68i and T300, which enables them to enlarge the text on the screen. Mobile magnifiers are screen magnification software for mobile phones which enlarge and enhance all items on the mobile phone display. They automatically detect and magnify the area of interest as the user navigates across the screen. Currently mobile phone magnifiers are produced by Rammland and Tieman and can be used on a number of Nokia phones and the Siemans SX1. Both magnifiers provide an auto zoom function, the ability to zoom onto different areas of the display and user selectable colour schemes. They both also have keyboard short cuts or hot keys. The Rammland magnifier magnifies the screen up to 6× and can provide an optional border with selectable size and colour to make it easier to identify the magnification window. Both magnifiers are quite expensive. Rammland provides the option of a free subscriber identity module (SIM) or a pay as you go handset on any network with a mobile phone magnifier already installed.

Orange offers an automated personal assistant service controlled by voice recognition. It can be used to store and retrieve numbers and as a personal organiser. It is still available to blind and visually impaired people, though it has been withdrawn from general use.

The Alva Mobile Phone organiser distributed by the RNID is a personal organiser with Braille input keys, Braille display and synthetic speech. The Alva has organiser functions, including calendar, diary, clock, basic word processing, calculator, contacts directory and a "to do" list. It also has a mobile phone and can be used in hand-free or headset mode. It is priced at the level of a Braille notetaker.

The OWASYS 22C is a talking mobile telephone produced by the Spanish company, OWASYS S. L. The telephone is shown in Figure 12.27, which is annotated to show the purpose of all the controls. All functions on the OWASYS mobile phone are spoken using synthetic speech. There is no visual display and the buttons are well spaced and easy to find. It can write, send and receive text messages. It has common mobile phone functions, such as caller identification and phone book, where up to 250 contacts can be stored. There is an embedded antenna, a vibrating alarm and a choice of 56 ring tones. It can be used on any network either on a contract or as pay as you go.

Talks is synthetic speech software which can be incorporated into mobile devices. It uses a Simbian 60 or 80 operating platform. This relatively new platform, has been adopted by several of the main manufacturers. Talks can be purchased as an upgrade to an existing phone or on a pay as you go basis. It is compatible with several Nokia models and Siemens SX1.

Mobile Accessibility Software from Rammland uses synthetic speech to allow users to access mobile phone functions. It works with the Simbian series 60 operating systems and provides simple menus that give speech access to the following:

- The number dialled, as well as received and missed calls.

- The number of the caller or their name, if it is in the phone book, as well as the date and time of the call.
- Text and MMS messages and the content of messages being sent, which is spoken while the user types.
- The address list, with each contact having 20 fields for information.
- User profiles, which allow the ringer tone, vibration alert, text alert and other features to be changed.
- The calendar, calculator, text notes and recorded voice notes for the user or to be sent *via* text.
- Sending and receiving emails, though some sighted assistance may be required to set up the system.
- Date, time, battery level and signal strength.

This software works with several Nokia models and the Siemens SX1.

Mobile Speak from Tieman is a screenreader package for mobile phones that allows access to most of the functionality of the handset. There are two levels of detail for new and advanced users respectively, as well as keyboard shortcuts. Words and/or characters can be announced using the configurable keyboard echo. The speech settings can be configured for speed, pitch, volume and three different punctuation variants. The speech synthesiser is included. There is also a sound recorder, a calculator, FExplorer to manage files and folders and two audio profiles, one of which enhances the sounds when earphones are used. The system works on several Nokia models, of which Nokia 7610 is recommended, as well as the Siemens SX1.

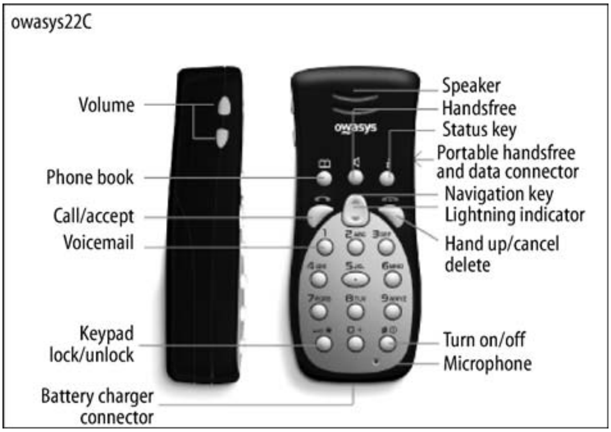


Figure 12.27. OWASYS 22C mobile phone (photograph provided and reproduced by kind permission of OWASYS S. L., Spain)

12.9 Chapter Summary

An increasing number of activities in modern society are information-based. The use of electronic media has completely changed the way most information is created, stored, retrieved and accessed. There are a number of different ways in which information can be categorised, including by the technologies used to present or transmit, store and manipulate the information, the applications of this information, and the assistive technologies required to make this information accessible to blind and visually impaired (or other disabled) people. The discussion in the chapter was based on the technologies used to present and transmit the information. This gives the three categories of printed information, computer- (and Internet-)based information and telecommunications-based information.

The assistive technologies for accessing printed information include low vision aids, audio reading systems and tactile transcription technologies. Dickinson's taxonomy of optical technologies was used to provide a framework to group the low vision assistive technologies. In the area of audio and tactile assistive technologies, both the innovative ideas of the READ-IT project and the traditional tactile methods of Braille were presented.

Access to computer information requires accessible computer systems. A generalised computer accessibility architecture was introduced (Figure 12.18) and used to discuss the accessibility options in the three main categories of system configuration, input and output. A further dimension is the sensory modality used to access information and the associated formats in which information can be presented, giving the three categories of visual, audio and tactile information.

The Internet plays an important role as an information resource and therefore disabled people require Internet access to ensure that they do not become the digital 'have-nots' on the wrong side of the digital divide. In addition, the Internet has an important role in providing access to services and communication that can increase the independence of disabled people. For instance, disabled people who find town centres and local shops inaccessible can do their shopping over the Internet, as long as it is designed to be accessible. The Internet also offers a range of communication options to disabled people. Guidelines for accessible website design and accessible web authoring tools (to enable disabled people to produce their own accessible websites) were presented in this chapter. However, it should be noted, that despite the benefits of the use of the Internet to provide information, there are currently a large number of people who do not have access, are unable or do not wish to use it. Therefore, it is important that the Internet supplements rather than supplants other means of providing information.

The final information category in the chapter was telecommunications, another area of rapid technological changes. The increasing popularity of mobile telephones and the use of text messaging has raised accessibility barriers for blind people. Mobile telephones are increasingly being used as an information portal, which could have great benefits to blind and visually impaired people, if the technology is made accessible (at a reasonable price). Current accessibility technology for mobile telephones is generally based on screen readers.

One of the aims of this review chapter was to provide an introduction, overview and context for the next four chapters (Chapters 13–16) of the book. In these chapters, a number of experts will give detailed presentations of the different assistive technologies that provide access to information. Screen magnifiers and readers are examined in Chapter 13 by Paul Blenkhorn and Gareth Evans of the University of Manchester. Rüdiger Hoffmann of the University of Dresden discusses the technology for speech, text and Braille conversion in Chapter 14. James Fruchtermann, who has been a prime mover in the DAISY Consortium, reviews the technologies for accessing book and documents in Chapter 15. Access to symbolic information is a particularly demanding technological task and in Chapter 16, David Crombie and Roger Lenoir of Dedicom, formerly the National Library for the Print Impaired in the Netherlands, explain how music can be made accessible to blind people.

Acknowledgement. The authors would like to acknowledge the kind assistance of the Humanware Group, New Zealand for providing several photographs of the Pocketviewer (Figure 12.6) and giving permission for their use in this book.

Thanks are due to Dr. Jacek Jelonek of the Department of Computer Science and Management of Poznan University of Technology, Poland for permission to reproduce various graphics and photographs (Figures 12.12, 12.13 and 12.14) from the Read IT project.

The authors would like to acknowledge Freedom Scientific, USA for providing several product photographs (Figures 12.1, 12.17, 12.21, 12.23), giving permission for their use in this book and directing the authors to information about these devices.

Thanks are also due to British Telecommunications plc, UK (Figure 12.24), Cobolt Systems Ltd., UK (Figure 12.25) Compusult Ltd., Canada (Figure 12.26), and OWASYS S.L., Spain (Figure 12.27) for the provision of photographs and permission to use them in this book.

Questions

Q.1 Describe the optical principles of the four types of low vision aid in Dickinson's taxonomy.

Q.2 List the advantages and disadvantages of the distributed and integrated constructions for closed circuit TV reading systems.

Q.3 Name two tactile alphabets and discuss their similarities and differences.

Q.4 Draw a diagram for an accessible computer system architecture. Give a brief description of each of the input and output devices in the architecture.

Q.5 Briefly describe the main accessibility options for visually impaired and blind people in
- (a) WINDOWS
- (b) LINUX

Q.6 Discuss the role of the different W3C Guidelines in ensuring website accessibility for visually impaired and blind people. Do you consider that any of the guidelines are not relevant to this group of end-users? Suggest additional guidelines that would make it easier for blind and visually impaired people to use websites.

Q.7 List the accessibility options available to visually impaired and blind users of:
- (a) A fixed line telephone
- (b) A mobile telephone

Projects

P.1 Dickinson (1998) states that “telescopes are a very effective way of producing magnification without changing the working space”, but they have a number of disadvantages that prevent their wide application.
- (a) Investigate what these disadvantages are.
- (b) Investigate how modern electronics and miniaturisation could offer a solution to these problems.
- (c) Draw up specifications for an improved design of magnification telescope.
- (d) Build your telescope.
- (e) Draw up and implement a programme for testing technical aspects of the telescope, including performance, reliability and safety.
- (f) If possible, obtain appropriate ethical permissions for testing the telescope with end-users and carry out tests with end-users. You should record end-user attitudes to the device and their suggestions for improvements, as well as their performance with the device. This will require an appropriate experimental design to ensure that the end-users are not subjected to any risk, as well as appropriate attention to data confidentiality, providing full information and consent issues.

P.2 Obtain a copy of the Read IT project (Chmiel *et al.* 2005) and carefully review its findings.
- (a) Assess the extent to which the technological performance obtained meets the objectives of the project.
- (b) Devise an enduser test programme for the Read IT system. You should take into account ethical, safety, confidentiality and consent issues and consult, for instance the British Psychological Society’s Ethics and Code of Conduct document.
- (c) Construct plans for:
  - (i) Commercialising the system or
  - (ii) Distributing the system to end-users using social mechanisms.

Discuss the advantages and disadvantages of the two approaches.

P.3 Braille keyboards are rarely used to provide computer input.
- (a) Draw up a survey questionnaire and accompanying information sheet for potential users to investigate their attitudes to Braille keyboards and how they could be made more popular.
- (b) Obtain appropriate ethical permissions and implement the questionnaire.
- (c) Analyse the results of the questionnaire and use them to produce a number of recommendations.

P.4 Design and build a portable Braille fingerspelling communication device for deafblind people. This will involve the following tasks:
   (a) Communication with a group of deafblind people on their requirements from the device. This will involve choosing an appropriate methodology, obtaining ethical permission and contacting appropriate organisations of or that work with deafblind people.
   (b) Using the results of task (a), draw up performance and other specifications for the device.
   (c) Construct a prototype device.
   (d) Devise and implement a test programme for the device to test its technical performance, including functionality and safety.
   (e) If possible, obtain the appropriate ethical permission and test the device with end-users. You should record end-user attitudes to the device and their suggestions for improvements, as well as their performance with the device. You should take account of all the issues listed in P.1.

References

ADOBE, 2005a, http://www.adobe.com/products/acrobat/access_faq.html

ADOBE, 2005b, http://www.adobe.com/enterprise/accessibility/pdfs/acro7_pg_ue.pdf

ATAG, undated, http://www.w3/TR/ATAG10

Bailey, I.L. 1982. The Honey Bee lens: a study of its field properties. Optom. Monthly, Vol. 70, pp. 275–278

Bettman, J.W., and McNair, G.S., 1939. A contact-lens telescopic system. Am. J. Optom. Physiol. Opt., Vol. 22, pp. 27–33

Chmiel, J., Stankiewicz, O., Switala, W., Tluczek, M., and Jelonek, J., 2005, Read IT Project Report: A portable text reading system for the blind people, Dept. of Computer Science and Management, Poznan University of Technology, Poznan, Poland

Dickinson, C., 1998, Low Vision: Principles and Practice, Butterworth-Heinemann, Oxford, U.K., ISBN 0 7506 2262 8

Feinbloom, W. 1977, Driving with bioptic telescopic spectacles (BTS). Am. J. Optom. Physiol. Opt., Vol. 54, pp. 35–42

Genensky, S.M., 1969, Some comments on a closed circuit TV system for the visually handicapped, Am. J. Optom. Arch. Am. Acad. Optom., Vol. 46, pp. 519–524

Genensky, S.M., Moshin, H.L., and Peterson, H.E., 1973, Performance of partially-sighted with Randsight I equipped with an X-Y platform, Am. J. Optom. Arch. Am. Acad. Optom., Vol. 50, pp. 782–800

Gregory, W., 1996, The Informability Manual, HMSO, London, U.K., ISBN 0 11 702 038 9

GNU, undated, http://www.linux.org/info/gnu.html

Hersh, M.A., and Johnson, M.A., 2005, Information technology, accessibility and deafblind people, proceedings of Association for Advancement of Assistive Technology in Europe Annual Conference, Lille, France

Hersh, M.A., and Johnson, M.A., 2006, Accessibilty of PDF documents, AXMEDIS Conference, Leeds, U.K.

Kelleher, D.K., 1979, Driving with low vision. J. Vis. Imp. Blind. Vol. 73, pp. 345–350

Lagarde, C., 2007, Personal communication

LINUX, undated, http://www.linux.org

Potts, A.M., Volk, D., and West, S.S., 1959, A television reader as a subnormal vision aid, Am. J. Opthalmol., Vol. 47, pp. 580–581

UWCAG, undated, http:www/w3.org/TR/UNDERSTANDING-WCAG20/

WB,-, http://www.watchfire.com/products/desktop/bobby/default.aspx

WebXACT, undated, http://webxact.watchfire.com/

Resources

The Pocketviewer manufacturer is Humanware Group for which the website is www.humanware.com. Distributors of the Pocketview products include Florida Reading and Vision Technology (www.floridareading.com), American Printing House for the Blind (www.aph.com) and RNIB (onlineshop.rnib.org.uk)

Report and Presentation material for the READ IT project can be viewed at the IEEE website www.computer.org/csidc under the 2004 student winners where the team from the Poznan University of Technology won second prize

A wide range of state of the art assistive technology products can be viewed at the Freedom Scientific website; www.freedomscientific.com

Compusult Limited of Mount Pearl, Newfoundland, Canada has an interesting range of assistive technology to make the workplace accessible; see the website www.compusult.com

Talking Caller Identification products http://www.talkingcallerid.com

13 Screen Readers and Screen Magnifiers

Learning Objectives

The use of personal computers and access to the Internet are rapidly becoming essential features of everyday life for all in the community. Whether the user is monitoring their bank account, sending an e-mail to a distant friend or working with an applications program like a word processing package, visually impaired people require assistive software to support them. This chapter discusses the design and use of common assistive software for visually impaired people. First, the visual screen interface can be enhanced using screen magnifier technologies to assist people with some degree of functional vision and, second, for those with no vision, screen readers provide an alternative interface presenting information *via* synthesised speech or Braille. Thus the learning objectives for this chapter include:

- Understanding the key functions of screen magnifiers and screen readers and how these support visually impaired users.
- Understanding the design options and architectural choices available to the designers of screen magnifiers and screen readers by the presentation of the historical development of these systems and a description of modern, commercial examples.
- Appreciating how self-magnifying and self-voicing applications differ from screen magnifiers and screen readers.

13.1 Introduction

Modern personal computers generally provide a multi-windowed, mouse-driven graphical user interface (GUI) that cannot be directly accessed by people who have visual impairments. Software tools exist that allow visually impaired people to interact effectively with computer systems. People with no functional vision (*i.e.* blind people) use screen readers. People with some functional vision (often referred to as partially sighted people) use screen magnifiers.

A screen reader presents the visual output of the screen to the user in synthesised speech or in Braille (by means of a refreshable Braille line). The user interacts with the computer system using the keyboard.

A screen magnifier presents an enlarged portion of the screen on a user's monitor. The user generally interacts with the computer in the normal way (*i.e.* using mouse and keyboard). Because only a portion of the screen is visible at any time, movement of the mouse, not only moves the mouse pointer, but can also change the portion of the screen that is visible to the user. Some users, with limited functional vision, can be supported by systems that combine a screen magnifier and screen magnifier. We will refer to such systems as hybrid magnifier readers and discuss their special characteristics towards the end of the chapter.

Screen readers, screen magnifiers and their hybrid combination are designed to make the operating system and most applications that run on that operating system accessible to people with visual impairments. As we will see later, these accessibility tools are quite complex and highly operating system specific. Moreover, accessing a complex operating system and application can be quite difficult for people who are 'technologically frail' because they not only have to access a complex computing environment but they also have to learn to use a complex accessibility tool—this is especially true for screen reader users. An alternative approach to providing accessible applications is to design or adapt applications so that the accessibility features (magnification or reading) are provided in the application itself. This leads to a set of applications that can be referred to as *self-magnifying* or *self-voicing*. The former are quite rare—although many applications exist that allow the font size and colour combinations to be adjusted to suit a user's needs. The latter are reasonably common and self-voicing word processors, web browsers and other tools exist. Whilst such applications are not the main focus of this chapter, we will deal with these briefly before its end. In addition, we will also consider a class of applications that are designed to make the use of standard applications easier to use with a screen reader—we will refer to these as application adaptors.

13.2 Overview of Chapter

First, we need to introduce some general concepts and terminology of GUIs. This information is relevant for the sections on both screen magnifiers and screen readers, since the issues are common to both.

There are then two main sections; the first discusses screen magnifiers and the second screen readers. The sections have a similar structure. Initially we focus on the user requirements and the user experience and identify the features provided by modern, commercial implementations of these tools. From a technological perspective, the implementation of both types of tool is quite complex. We will address these issues by giving a brief history of DOS and Windows screen magnifiers and screen readers before presenting the architecture of modern, commercial systems for Microsoft Windows. This focus on Microsoft Windows is justified because the vast majority of screen magnifiers and screen readers are designed for this platform. We then briefly consider the use and operation of hybrid screen magnifier readers.

To conclude the chapter, we briefly examine self-voicing applications and the application adapters designed to make screen readers easier to use.

13.3 Interacting with a Graphical User Interface

The goal of both a screen reader and screen magnifier is to provide access to a computer system so that a visually impaired user can access all the features of the operating system and most, if not all, of the applications present on the computer system. Before we discuss screen magnifiers and screen readers in detail, we need to present some of the key features of GUIs.

It is assumed that you have used a windowed GUI and are comfortable with the concept of a window. If you are, you will be aware that, whilst there may be many windows open on the screen, only one window has the focus at any given time. You can verify this by opening a number of copies of a text editor or word processor on a computer and typing at the keyboard. The keystrokes are routed to one window—the window with the *focus*. You will be able to change the window that has the focus by clicking on it with the mouse or using a keyboard command (Alt-Tab in Windows).

We need to refine further the concept of *focus* to a finer level of granularity, *i.e.* to say that within the window that has the focus, the focus is on a particular element of the window. As I type this chapter in Microsoft Word, the focus is currently in the client area of Word (*i.e.* the part of Word into which text is typed). I can change the focus by using the mouse—for example, I can move the mouse pointer to the File menu item and, after I have clicked, the current focus becomes the File menu. I could change the focus by using a Word keyboard shortcut. For example, Alt-F will activate the file menu, causing the focus to move to this menu. When the File menu appears using the shortcut, the first menu item is highlighted (in the case of my system the 'New' menu item), so, not only does the menu have focus, but a single item within the menu has the focus, in this case the 'New' menu item. I can change the focus by using the arrow keys to scroll up and down through the menu. If an item has a submenu (for example, in the Word File menu the Send To menu item) then I can change the focus to the submenu by selecting the right cursor key. I can then scroll up and down this submenu. As a further example, I can activate the spell and grammar checker of word (either by using the mouse or by using the keyboard shortcut F7). This causes a pop-up *form* to appear which takes the focus. By a *form*, we mean a window that has a number of *controls* on it. The *controls* are items such as buttons, text list boxes, radio buttons, *etc*. When the form appears, one control within the form has the focus (this is indicated by visual highlighting); in the case of the spell and grammar checker, this is the text list box that has the first suggestion for changing the spelling. I can change the focus within the form by hitting the Tab key, which will move the focus from control to control. The important point is that, at any one time, one element has the focus and that the user, the application and the operating system are capable of causing a change of this focus. The application and the operating system may change the focus when a pop-up menu box appears to signal an error, for example.

Screen Magnifiers and Screen Readers both need to keep track of the focus so that relevant information is presented to their users. This might appear quite straightforward at first glance but it can become quite complex, especially in the case of a screen magnifier in which the user controls the mouse as well as the keyboard. For example, as I type this document into my word processor, the focus is quite clear; it is the flashing cursor that indicates the current point of text entry. However, when I move my mouse over the client area of the word processor, it becomes the caret that shows the position of the mouse. I can use my mouse (or keyboard shortcuts) to highlight a number of words, which may span over several lines, in which case the focus becomes the block of highlighted text. If the mouse moves outside the client area, it will become a pointer and allow me to select menu items. If I do so, the focus becomes the currently highlighted menu item. If I select a menu item that causes a pop-up form to appear, the focus falls on the item that has the focus in the pop-up form, which will generally be displayed in the middle of the screen. The point to note is that the focus changes, not only in its position on the screen, but in the type of information that has the focus and the way in which the focus is indicated. For a sighted user, the focus is presented by visual means. For example, the cursor in a word processor flashes, menu items are presented in different colours, and buttons are highlighted by subtle changes in their outline. Accessibility tools must keep track of the focus and determine what to display, the techniques by which they do so are discussed in the sections that follows.

Before we finish with the topic of focus, we should identify one further, quite complex example. This concerns the use of a spreadsheet (we will confine our discussion to Microsoft Excel, but it should apply to most spreadsheets). In the client area of a spreadsheet, the focus is generally on the current cell and the user can move from cell using the arrow keys. The current focus is visually indicated by visual highlighting of the cell and by highlighting the row and column that contain the cell (see Figure 13.1, which shows cell B3 as having the focus).

However, when a cell has the focus, not all information is accessible from that point. If the information in the cell is derived from a formula (as it is in Figure 13.1), the calculation for the formula is presented in the *fx* text box. The user of a screen magnifier will want to be able to see this information at the same time as they view the contents of a cell. The user of a screen reader will need to be informed that the data presented is the result of the application of a formula and what that formula is.

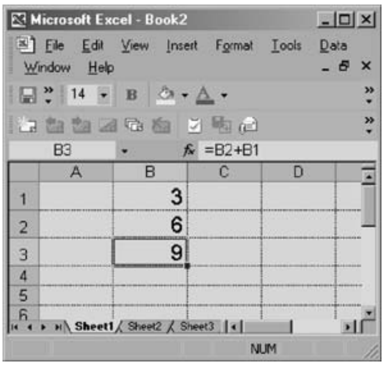


Figure 13.1. Example of using a spreadsheet

In summary, the user of both a screen reader and screen magnifier must always be given access to relevant information. This is generally found at the focus and screen readers and screen magnifiers must track the focus.

13.4 Screen Magnifiers

13.4.1 Overview

The purpose of a screen magnifier is to enlarge a portion of the visual display of a computer system so that it becomes visible to person with visual impairments. As discussed below, the whole of the computer screen may be taken over with the enlarged image (full screen magnification) or part of the screen may be used to present the enlarged image with the rest of the screen being displayed at normal resolution (part screen magnification, of which there are many types). In both cases it is normal for the enlarged image to be centred on the current focus. So, if the user moves the focus, the magnified portion of the screen should change. If a pop-up window appears, the enlarged image will be centred on the element of the pop-up window that has the current focus.

Commercial screen magnifiers generally provide enlargement factors in integer multiples from 2× to 16× (some magnifiers support 32× magnification). Some magnifiers, for example Ai Squared's BigShot (AiSqBi, undated), provide more modest levels of magnification, between 1.05× and 2×. In general usage, magnification levels are the same for the horizontal (x) and vertical (y) directions. However, some screen magnifiers allow the x and the y direction to be controlled independently of one another. So it is possible, for example, to have an image that is enlarged 4× in the y direction and 3× in the x direction.

It is important to realise that even with full screen magnification and relatively the modest enlargement factors, the user sees a much reduced portion of the screen. The proportion of the original image that can be seen by the user at any one time is

$$\frac{1}{x\text{Factor} \times y\text{Factor}}$$

where xFactor and yFactor are the x and y direction enlargement factors respectively.

Even at modest levels of magnification (say, 2× in both directions) and full screen magnification, the user can only see a quarter of the screen at any one time. When high magnification levels are used, the proportion of the original image that can be seen is very small—at 16× enlargement, it is 1/256th of the original image. This can pose problems for the users of screen magnifiers in that orientation (where am I?) and navigation (where can I find?) become difficult.

Another important issue concerned with magnification is that when magnified, text can appear quite *blocky*. In the example shown in Figure 13.2, note especially the blockiness of the *w*, the *a* and the *g*. This blockiness is referred to as *aliasing*

Figure 13.2. Text magnified at 8× showing aliasing

and most screen magnifiers employ smoothing algorithms to reduce the effects of aliasing. We shall discuss this issue in more detail later.

Before considering more detail of the operation of screen magnifiers we should present some terminology that we will use consistently throughout the rest of this section (the above uses terms quite loosely).

We will refer to:

- The image displayed on the computer screen before magnification (*i.e.* the screen as presented by the operating system) as the *original image*.
- The result of magnification (whether this takes over the full screen or part of it) as the *enlarged image*.
- The multiplicative factors used to enlarge the image as the *enlargement factors* and we will assume that, unless we state otherwise, the enlargement factor is common to the x and y directions, so that, if the enlargement factor is 2×, the enlarged image is doubled in both width and height.
- The different ways in which the enlarged image can be presented to the user are called *magnification modes*. These include full-screen magnification where the whole display is used to present the enlarged image and the set of part-screen magnification modes where the enlarged image is presented in one part of the screen and the original image can also be seen elsewhere.
- The blockiness of text and other parts of the enlarged image is referred to as *aliasing* and the algorithms used to reduce aliasing as *smoothing algorithms*.

13.4.2 Magnification Modes

Most screen magnifiers provide users with a number of different magnification modes. These are described below. Where possible each of the magnification modes is illustrated by a screen shot of the mode operating on a computer system. It should be noted that for technical reasons, which will become apparent when the architecture of screen magnifiers is discussed later, it is not possible to take screen shots using the standard Windows method (using the Print Screen button). Consequently, images have been taken by photographing a screen, this means that the quality of the image presented is somewhat reduced, but they should be sufficient for illustrative purposes.

13.4.2.1 Full Screen Magnification

As its name suggests, the entire screen is filled with the enlarged image as shown in the example of Figure 13.3.

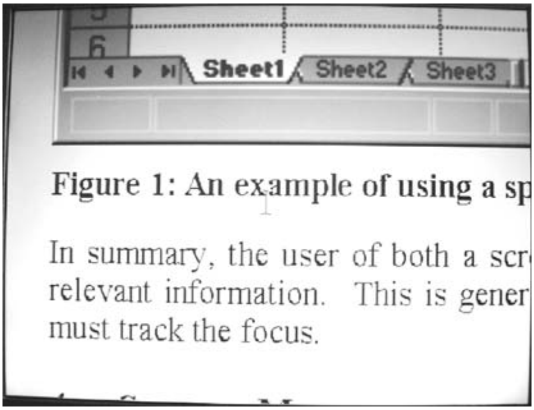


Figure 13.3. Full screen magnification (6× magnification)

This is a very popular mode, which a large number of users seem to prefer. It has the advantage of maximising amount of information that is presented to the user. The disadvantage of full screen magnification is that users may find it difficult to keep track of where they are in relation to the full screen and the boundaries of the application window. For example, if the user wishes to reach the Start menu on a Windows machine they will not necessarily know how far they have to move the mouse to find it. This is a relatively trivial example, because the position of the Start menu is usually fixed to the bottom left hand corner and the user can simply employ a strategy of moving the mouse until the bottom left hand boundary of the enlarged image is reached. Alternatively, the user may use a keyboard short cut to activate the Start menu. The problem is more acute when the desired location is not fixed in relation to the boundaries of the screen. For example, when the user is scanning for information in a document or web page or when they wish to find a particular point in a window. Some of the magnification modes identified below address this issue to some extent, at the expense of reducing the size of the enlarged image and hence the portion of the screen that is enlarged. The extent to which these modes address this problem and, indeed, introduce problems of their own is explored below.

13.4.2.2 Lens Mode

In this mode, the user is provided with a simulated lens that displays the enlarged image and which can be moved around the original image. The original image is visible outside the boundary of the enlarged image as can be seen in the example of Figure 13.4. It is analogous to reading a paper document with a magnifying glass.

In general, the size of the enlarged image can be controlled so that it takes up a greater or lesser proportion of the screen. The disadvantage of this mode, and indeed most of the other part-screen magnification modes discussed below, is that less information can be presented in magnified form to the user. This reduces the amount of *local contextual information* that a user can view at any one time. By

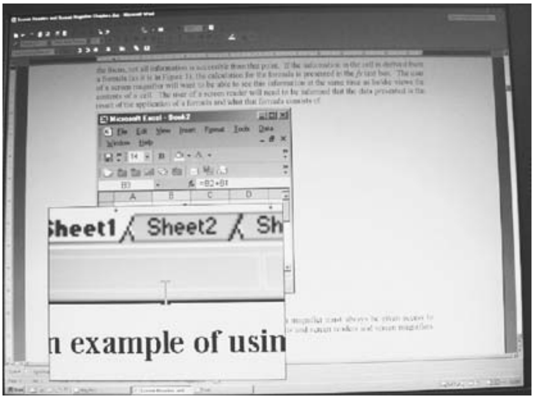


Figure 13.4. Lens magnification (6× magnification)

local contextual information we mean information adjacent to the focus. This is an important issue when using a word processor with a magnifier. As the enlarged image is reduced in size, fewer words are visible on the screen at any one time. This may make reading difficult, especially at very high magnifications when only a part of a word may be visible.

As noted earlier, even at a magnification level of 2×, only a quarter of the screen is visible for a full screen magnifier. If a lens magnifier is used that occupies 1/16th of the screen (a typical default size for a lens magnifier), only 1/64th of the original image is magnified. The possible advantage of a lens magnifier is that the user has more information about which part of the screen is being magnified. If the user's functional vision is sufficient to recognise relatively coarse features in the original image, they might be able to recognise window boundaries and hence be able to move to particular areas by sight. Even if the user cannot recognise any features in the original image, they should still be able to see where the magnified portion is in relation to the full screen. This makes finding fixed locations—for example, the Start menu—rather easier.

In short, there is a trade-off between size of the enlarged image against and understanding of which part of the original image is being displayed. The choice of which mode to use (full, lens or the more esoteric modes that follow) depends on the preferences of the user, the extent of their functional vision and the strategies that the user follows to interact with their computer.

13.4.2.3 Fixed Area Magnification

In this mode, part of the screen is presented as an enlarged image and part of the screen is presented as the original image; a 'split-screen' presentation. In this section we will describe this magnification mode with one part of the screen presenting the original image and one part of the enlarged image. It should be noted, however, that this mode is often used with a CCTV-based document enlarger. The output from this device—an enlargement of a paper document—is presented

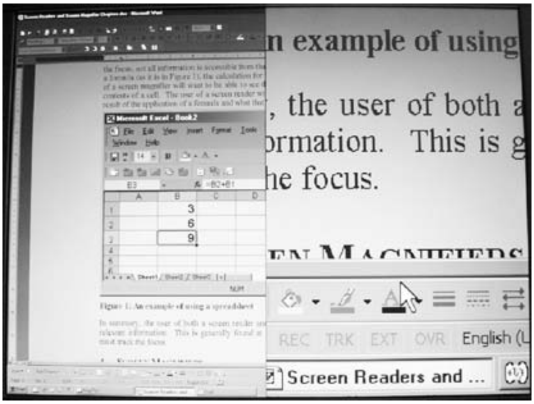


Figure 13.5. Fixed area magnification. Shows the enlarged image to the *right* of the screen (6× magnification)

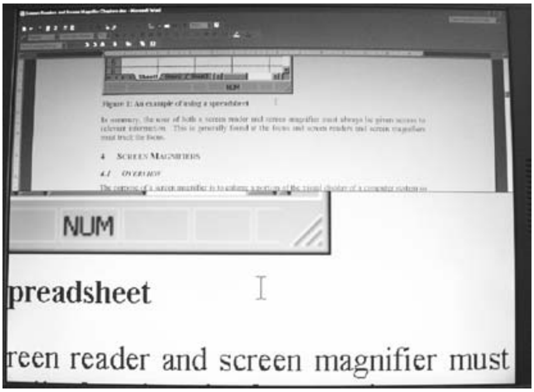


Figure 13.6. Fixed area magnification. Shows enlarged image at *bottom* of screen (6× magnification)

in the portion of the screen that would be occupied by the original image. This style of presentation allows a user to view enlargements of paper documents and the computer screen at the same time and on the same screen. We will not consider this style of operation further and focus on magnifiers that present the enlarged and original image side-by-side.

Typically, the area that presents the enlarged image is quite large (often as big as half the screen area) and whilst the user can decide where to place it relative to the original image, the user will rarely move the position of the enlarged image as they use the system. Generally, the enlarged image is positioned along one edge of the screen. In vertical configurations it occupies the left or right hand side of the screen (see Figure 13.5) and in horizontal configurations the bottom or top of the screen (see Figure 13.6). However, some systems allow the enlarged image to be presented in any part of the screen.

There are a number of issues to consider when fixed area magnification is used because different magnifiers work in different way. The first issue concerns what is displayed in the portion of the screen given over to the original image. One approach is to compress the original image so that it fits its new boundaries. Whilst this causes some degree of distortion to the original image, if the user has sufficient functional vision, they may be able to make out the position of the mouse pointer and other salient features such as window boundaries. This may help address the orientation problems associated with full screen magnification. Another approach is to display that portion of the original unmagnified image that would normally occupy that screen in the area that is unmagnified. This approach is shown in Figures 13.5 and 13.6.

If the magnifier uses the approach where only a part of the original image is displayed, a second issue arises. There are two concerns, both of which are prompted by the question of what happens when the focus moves to part of the original image that is obscured by the enlarged image? To discuss these concerns consider the display shown in Figure 13.6 where the original image is displayed in the top half of the screen and the enlarged image is displayed in the bottom half of the screen. As shown in this figure, the enlarged image corresponds to a magnified portion of the original image. Two mouse pointers are visible (shown as a caret in Figure 13.6), a magnified one in the enlarged image and the original unmagnified mouse pointer in the original image. Suppose that the mouse pointer in the original image is now moved so that it is moved into the area occupied by the enlarged image, *i.e.* the mouse pointer moves to a portion of the original image that is covered by the enlarged image. There are two questions. First, what should be displayed in the original image? Second, what should be displayed in the enlarged image?

Regarding the original image, one option is to scroll the original image so that the portion of the original image displayed corresponds to that shown in the enlarged image. From a technical point of view, this is non trivial, but it does have the advantage that what is displayed in the original image reflects what is presented in enlarged form in the enlarged image. The other approach is to leave the original image unchanged so that, in the example we are considering, the top half of the screen remains unchanged, but the bottom half magnifies information held in the original image that cannot be seen on the screen. This is much easier to do. It may well be acceptable to the user, especially as the user will be able to reverse the positions of the magnified and original images if desired, in the case of our example by making the top half of the image the enlarged image and the bottom half the original image. Some screen magnifiers support this automatically. So, continuing our example, if the mouse pointer was moved from the top half of the screen (original image) into the bottom half of the screen (enlarged image), the top half would become the enlarged image providing an enlargement of the bottom half of the screen, which then shows the original image. This may be a bit disconcerting for the user, but if they are prepared for this mode of operation, it may well be acceptable.

The second issue concerns the information that is displayed in the enlarged image when the focus occupies part of the original image that is obscured by the enlarged image. As noted above, one option is for the enlarged image centred on the system's focus to be presented. This seems the most desirable way to operate.

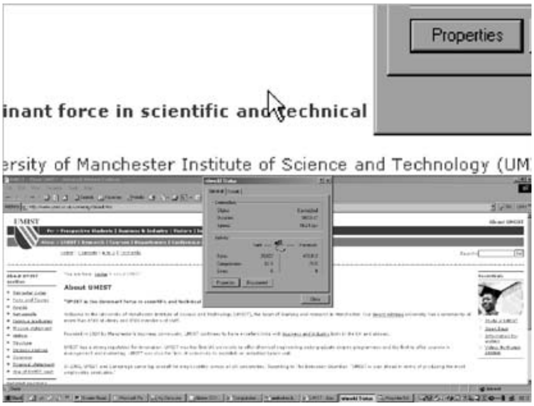


Figure 13.7. Microsoft Magnifier. Shows the enlarged image docked at top of screen. Note that the screen capture mechanism used has suppressed the mouse pointer in the original image—this would appear in the same relative position as that shown in the enlarged image

However, some magnifiers do not allow the portion of the original image that is under the enlarged image to be enlarged.

It is worth noting at this point that the Microsoft Magnifier that is supplied as part of Windows XP and Windows 2000 may be thought of as a fixed area magnifier. Project P.1 at the end of the chapter explores the features of this magnifier in more detail, but it is worth briefly describing its features here because most readers will have access to it. The Microsoft Magnifier is a fixed area magnifier that can be positioned anywhere on the screen. The size of the magnified portion can by easily changed, but the area that it occupies can never be greater than half the screen size. The magnifier can be docked to any of the four screen edges. When docked, the rest of the screen (with the exception of the system tray) is resized so that all of the original image can be seen. The original image is not distorted because applications are resized and icons repositioned to take advantage of the space available for the original image (See Figure 13.7). So, in this case, the enlarged image is presented along aside, rather than on top of, the original image.

The Microsoft Magnifier does not have to be docked, but can occupy any position in the screen. When it is not docked, the original image is shown as normally displayed and the enlarged image corresponds to the current focus (see Figure 13.8). However, the magnifier will not magnify any portion of the original image that falls under the enlarged image—this is shown in the top right hand corner of the magnifier window in Figure 13.8.

13.4.2.4 Other Magnification Modes

Most commercial screen magnifiers provide the magnification modes described above. In this section we discuss other, more esoteric magnification modes that are provided by some, but by no means all, of the commercial screen magnifiers.

Figure 13.8. Microsoft Magnifier when not docked. Shows the cursor outside the enlarged image so that the capture process suppresses mouse pointer in original image. Note the *upper right part* of the enlarged image corresponds to part of the image that falls underneath the enlarged image and is therefore blank

13.4.2.4.1 Strip Magnification

This mode is similar to lens mode, in that the enlarged image corresponds to a small portion of the original image centred on the current focus. In strip mode, the enlarged image is a horizontal strip that extends for the full width of the screen—see Figure 13.9.

The idea is that this mode supports the reading of text. It is thought to be particularly effective when, the vertical (y direction) enlargement factor is set to be greater than that of the horizontal (x direction) enlargement factor. This has the effect of magnifying text to a level where it can be read, whilst, at the same time, maximizing the amount of material that is visible on the screen at one time (*i.e.* maximising the local context). This mode is thought to be particularly effective for users with more modest visual impairments, who require only a small amount

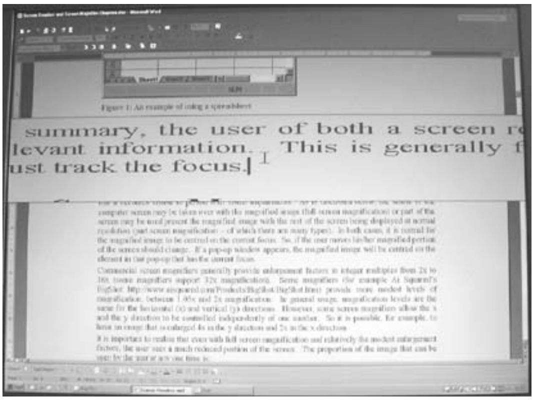


Figure 13.9. Strip magnification, horizontal enlargement factor = 2, vertical enlargement factor = 4

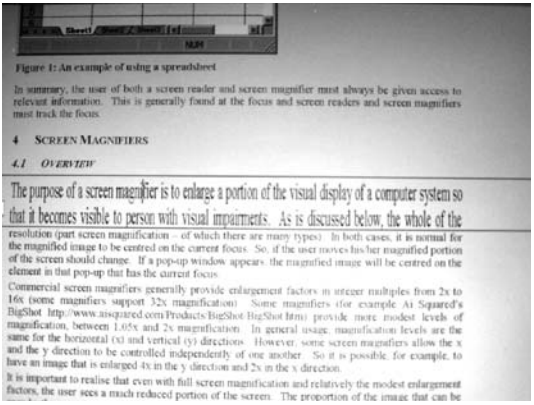


Figure 13.10. Strip magnification, horizontal enlargement factor = 1, vertical enlargement factor = 2

assistance to read text, when the y direction enlargement factor is set to $2\times$ or $3\times$ and there is no enlargement in the x direction. This has the effect of enlarging text (as the cost of some distortion), whilst presenting the full width of the screen in the enlarged image. Figure 13.10 shows a y direction enlargement factor of $2\times$ and an x direction enlargement factor of $1\times$.

The height of the strip can be controlled by the user to determine the number of text lines that are to be viewed. The Magnus magnifier (MagM, undated) devised by Sensory Software Ltd provides two versions of this mode. One where the enlargement image corresponds directly to the focus (*i.e.* the enlarged image is presented over the corresponding part of the original image—this is shown in Figure 13.10). In the second version, the enlarged image is offset from the corresponding part of the original image, so that a user with relatively good functional vision can view the same point in the original and enlarged images—see Figure 13.11.

It is not clear how widely strip magnification is used by visually impaired users. We would suggest that it would be of use to some users with relatively good functional vision. The mode has been widely used by people with some Specific Learning Disabilities (for example some dyslexic users), where the horizontal strip assists users in focusing on text. For some users, the original image outside the strip is obscured to assist them in focusing on the selected text.

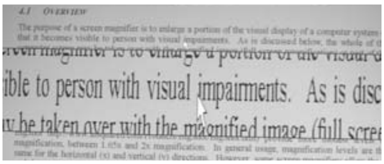


Figure 13.11. Strip magnification with offset. This shows horizontal enlargement factor = 2 and vertical enlargement factor = 4. Note that mouse pointer appears in the original and enlarged images

13.4.2.4.2 Multiple Windows

The modes discussed above give the user a single magnified window (which may be the full screen). There are circumstances where a user wishes to magnify more than one area of the screen. For example, a user may wish to work in one application, whilst keeping an eye on the system's status bar. This can be achieved in a multiple window system, in which the user can define multiple windows. Windows may be static (*i.e.* they magnify a fixed area of the screen, they do not track focus) or dynamic (they track focus).

A good example of where this mode could be used is when a user interacts with a spreadsheet (for example Microsoft Excel). As the user moves from cell to cell (using the cursor keys), not only is the client data updated, but also the highlighting of the row and column headers and the formula (*fx*) area are also updated (see Figure 13.1). If the user is using full screen magnification, it is likely that they will not be able to see the row and column headers and the formula area at the same time as viewing a cell—there is not enough room in the enlarged image. However, if the user could define three additional windows (one for each of the formula area, the column header and the row header) and these could be presented in the same enlarged image as the cell, the user can view all the relevant information on the same screen. The user will not have to leave the cell that they are working in and search for the formula area. This should greatly increase user efficiency. The formula area can be handled in a different way to the row and column headers. The interesting part of the formula area is fixed on the screen and this should be made a static area (one that does not change its view with focus changes). The row and column headers should be made active areas, which track changes in focus so that the highlighted row/column number/letter is always presented. This style of interface should greatly aid a user when they need to view information from different parts of the original image at the same time. However, it is not a trivial operation for a visually impaired user to configure such as system—determining where to place the windows and what characteristics to give them can be difficult. It is suggested that this style of interface is best used by people who use complex applications on a regular basis where configurations for particular applications can be stored and recalled.

13.4.3 Other Interface Considerations

In addition to the screen magnification modes, there are a small number of interface considerations that need to be presented; these are:

- Behaviour of the mouse pointer in the enlarged area.
- Smoothing of text.
- Colour inversion.
- Controlling the magnifier through other means than the mouse and keyboard.

13.4.3.1 Mouse Pointer Behaviour

A screen magnifier will track the mouse pointer as it moves and will place the mouse pointer roughly in the centre of the magnified area. Suppose that full screen magnification is being used and the user moves the mouse pointer a short distance. The designer of the screen magnifier has two options. First, the enlarged image can be moved so that the pointer is always at the centre of the screen. Of course, at the edges of the original image, the pointer cannot be centralised in the enlarged image because this would mean that the enlarged image would contain an area that was not present in the original image. In this style of behaviour, every movement of the mouse (or indeed anything else that causes the focus to change—for example typing text) will cause the image to shift (see Figure 13.12).

The second approach is to move the mouse pointer in the enlarged image and not to shift the enlarged image when the mouse movement is small, *i.e.* it does not move close to the edge of the enlarged image. In this case, the enlarged image remains static and the mouse pointer moves (see Figure 13.13).

In our experience, users prefer the second type of behaviour—*i.e.* scrolling the enlarged image only when the mouse pointer moves close to the enlarged image's edge. The reason for this is that in the first style of interaction, the enlarged image is moves whenever the mouse pointer moves. This means that the user is presented with a constantly changing image, which can prove very tiring. Allowing the enlarged image to remain constant when mouse pointer movements and other changes in focus are relatively small (*i.e.* well within the enlarged image) means

a

b

Figure 13.12a,b. Example of mouse movement. This shows the mouse pointer in **a** has been moved to the *right* in **b**. The display has scrolled to the right. The mouse pointer has centralised the image during the movement

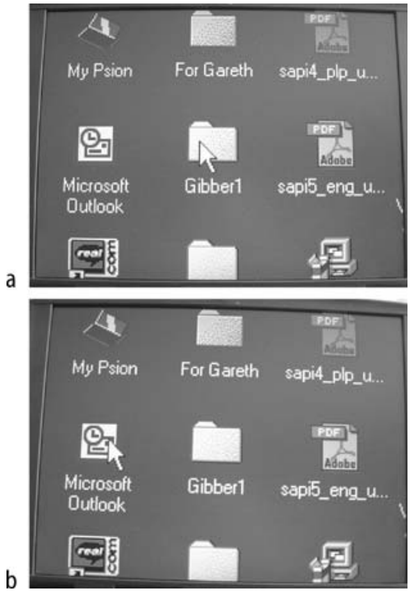


Figure 13.13a,b. Second example of mouse movement. The mouse pointer in **a** has been moved to the left. It can be seen in **b** that the mouse pointer has moved rather than the image (contrast with Figure 13.12)

that the user is presented with a much more stable image that is less tiring to use. Most commercial magnifiers use this style of interaction, but Microsoft Magnifier does not (see the project P.1).

13.4.3.2 Smoothing

As noted earlier, when text is enlarged, it can appear *blocky*; see Figure 13.2. Many users prefer to read text when it is smoothed and most screen magnifiers provide an option for smoothing. An example of unsmoothed and smoothed text is presented in Figure 13.14.

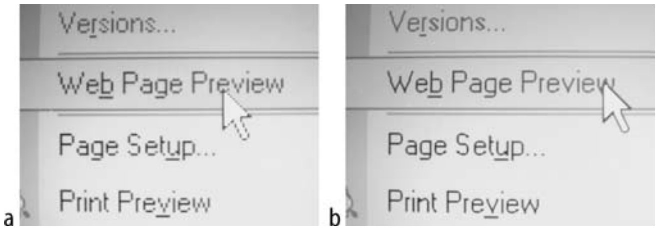


Figure 13.14a,b. Example of smoothing. **a** shows no smoothing (note the 'w' and the mouse pointer). **b** shows smoothing (note the slightly 'odd' formation to some characters). Enlargement level 6 $\times$

You should note the reduction in aliasing and the relative smoothness of the text. You will also perhaps note that the smoothed text is not as clear as if the original text had been presented in a larger font (rather than magnified and then smoothed) and that some artefacts of the smoothing process make some aspects of the characters look a little unusual, *e.g.* the *v* and the *w* in Figure 13.14. The reason for this is that the algorithms that carry out smoothing operate directly on a bitmap representation of the original image—they are image processing techniques—and consequently do not have access to the information that indicates which character is being displayed and thus cannot simply replace a character with its equivalent in a larger font. If screen magnifiers used some of the techniques employed by screen readers (see later), the identity of characters could be obtained. However, it is unlikely that simply enlarging the font size of any text elements would give an entirely faithful representation of the screen and would not address the enlargement of elements of the screen that were not text so this approach is not used by screen magnifiers.

Few details of smoothing algorithms for screen magnifiers are published—the only instance known to the authors is that in (Blenkhorn *et al.* 2002). Most algorithms have to be able to distinguish between the foreground colour (*i.e.* the text colour) and the background colour and there are often techniques for user intervention to support this. Users must be able to turn smoothing on and off. Smoothing is aimed at tidying up the presentation of text and can sometimes distort images.

13.4.3.3 Colour Inversion

Most magnifiers (including Microsoft Magnifier—see the project P.1) provide the user with the option to invert colours. This is provided because some visually impaired people have conditions where white text on a black background is preferred to the more prevalent representation of black text on a white background. Colour inversion works well on a monochrome image (see Figure 13.15) where black becomes white and *vice versa*. However, it produces different results—which may be less useful to users—on colour images, where for example blue can become yellow or peach. Figure 13.16 shows an example of inversion on coloured images. You can explore the effect yourself using Microsoft Magnifier (see the project P.1).

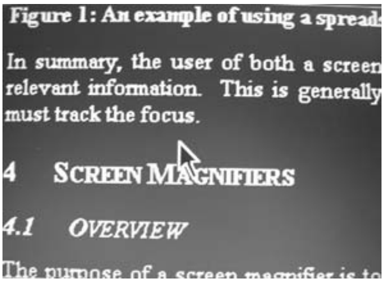


Figure 13.15. Colour inversion monochrome

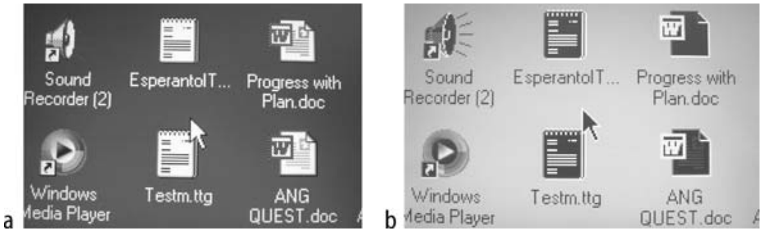


Figure 13.16a,b. Colour inversion of a coloured image. **a** shows original, **b** shows inverted

13.4.3.4 Controlling the Magnifier by Means Other than the Mouse and Keyboard

So far, we have assumed that a user will control the magnifier by using a mouse and keyboard in the standard way. Some alternative approaches to controlling the enlarged area have been used.

The most notable example is the probably the TASO system produced by Frank Audiodata (FraAudio, undated). This system provides horizontal and vertical sliders that control position of the user on the screen and also provide audio feedback indicative of the user's position. The system is also used with screen reader users. The benefit for screen magnifier users would appear to be that the user can determine the absolute position of the enlarged image relative to the original image by noting the position of the sliders.

There has been some investigation into the control of a magnifier through a joystick (Kurniawan *et al.* 2003). The motivation for this work is that users (especially older users who may be computer novices) often lose orientation and find navigation difficult driving a screen magnifier with a mouse. The idea is that using a joystick as an absolute positioning device (*i.e.* joystick position directly corresponds to the portion of the enlarged area) or as a mouse-like relative positioning device (with explicit directional prompts) may ease these problems. Limited testing with a number of older, computer novices indicates that the use of a joystick may improve usability of magnifiers for some users.

13.4.4 The Architecture and Implementation of Screen Magnifiers

In this section, we present the architecture of a modern Windows screen magnifier and outline the implementation options that a designer faces. We also present a brief history of magnifiers for the DOS and Windows operating systems in the hope that this will make clear the choices that are faced by the designer of a system.

In considering the history, we will focus mainly on the issue of how the magnifier acquires the original image (so that it can create the enlarged image). We will also indicate some of the methods used to display the enlarged image, but not all possibilities are considered. We will deal with other issues such as focus tracking and cursor display, when we present the architecture of a modern magnifier. For the purposes of the historical survey, we will consider the requirements of magnifier to be that:

- The magnifier must track the user's current working position—the focus.
- It must present an enlarged version of the image roughly centred on the focus.

13.4.4.1 A Brief History of Screen Magnifiers

There are a number of approaches that can be used by magnifiers to access the original image. We will consider the following:

- Custom hardware solutions.
- Exploitation of the underlying hardware.
- Operating system-based magnification (GDI/DD hooking).
- Device driver replacement.
- Device driver chaining.
- Video card overlay.

The first two approaches are perhaps best considered in the context of a simple operating system (for example DOS). A simplified view of the system's architecture is presented in Figure 13.17.

The other approaches are best considered in the context of a modern multiple-windowed GUI. As noted earlier, we will confine our discussion to the Windows operating system. A simplified view of the system's architecture is given in Figure 13.18.

13.4.4.1.1 Early Approaches for DOS-based Machines

Custom hardware solutions The earliest magnifiers were large character terminals that were effectively custom terminals for mini- and mainframe computers. With the advent of the personal computer, magnification was first achieved by adding custom hardware. This approach was used in the earliest magnifiers and some later systems (*e.g.* those produced by Frank Audiodata). The approach requires the construction of additional hardware. There were two variations to the approach. The first was to replace the computer's video card with a custom video card that

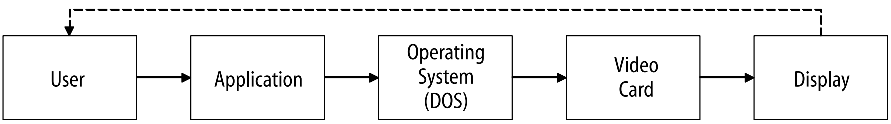


Figure 13.17. Simplified architecture of a typical DOS-based machine

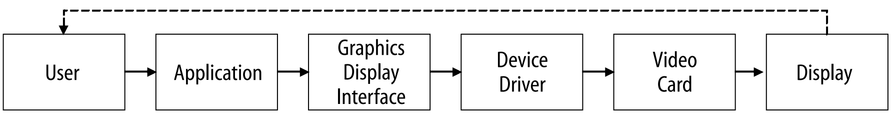


Figure 13.18. The simplified architectural view of a Windows-based machine

supported magnification (see Figure 13.19). The second method was to introduce custom hardware between the output of a standard video card and the display; the custom hardware enlarged the image and controlled the portion of the image displayed on the screen (see Figure 13.20).

An important issue with such systems was how to track the focus. In these early text-only systems, the focus corresponds to the text cursor. The approach adopted by VTEK DP series was not to track the focus but to provide the user with a joystick to control what portion of the image was displayed on the monitor (Griffith 1987). This approach was a hardware-only solution and in the days of less powerful computers with limited memory represented a reasonable solution. Another approach was to have memory-resident software that augmented the hardware magnification with tracking of the text input cursor position. This approach was adopted by TSI, amongst others, for their Vista system (Spiry 1987). To our knowledge, these approaches are no longer used. With the vast increase in computational power and memory, it has become far more cost efficient to develop software-only magnifiers.

Exploiting the underlying hardware These approaches used memory resident software programs that exploited the fact a video card provided a set of video display modes that could be selected by writing data to their control registers where the different display modes presented a different number of text columns and rows on the screen. The basic trick was to allow the operating system to create an image for a particular display resolution (in the early days the display resolution was the number of lines of text that could be presented) but to set the video card to another resolution by writing to its registers in an appropriate way. There were two basic approaches. One was to set the display such that additional lines of characters became available for the magnifier to present an enlarged version of the focus. The other approach was to allow the operating system to create an image that was at

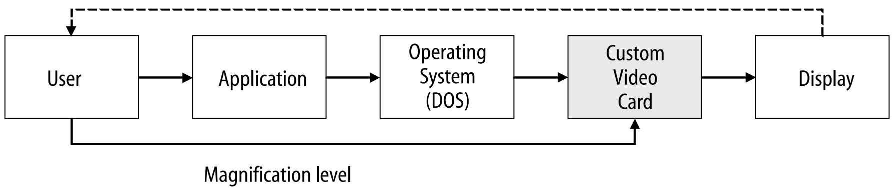


Figure 13.19. Hardware-based magnification using a custom video card. This replaces the standard video card and will enlarge a portion of the image under user control

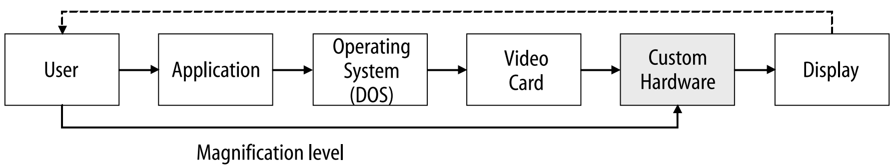


Figure 13.20. Hardware-based magnification using custom hardware. The custom hardware is positioned between a standard video card and the computer's display

higher resolution than required and then to present a portion of this image on a video card configured for a lower resolution.

The early versions of these systems were almost exclusively concerned with character enlargement and thus supported text but not graphics. It is thought that the first example of this approach was the Lyon Large Print Program (LLPP) (Gibson and Reiser 1987). Early versions of this system directly addressed the video registers of the control chip on the video card to produce a screen display of 32 lines rather 25 lines. The operating system was unaware of this change and produced its output on the top 25-lines of the display. Standard ASCII graphical characters, found in the upper portions of the ANSI character set, were placed in the bottom five lines of the video memory to simulate large characters. Sixteen enlarged characters could be displayed in the 'soft display' (see Figures 13.21 and 13.22).

This system was, therefore, a fixed area magnifier (see the earlier discussion on magnification modes). The HAL16 magnifier from Dolphin Systems used a very similar approach. Dolphin further developed this approach for use with EGA monitors. In this case, the expanded display rose to 43 lines, which enabled a larger 'soft window' to be produced, giving between 1 and 3 lines of magnified text depending on the magnification level set by the user.

Dolphin Systems developed two further variations on this approach. Hal40 changed the screen resolution mode to 40 × 25 characters, whilst operating system continued to function in 80 × 25 characters. Hal40 monitored the cursor position

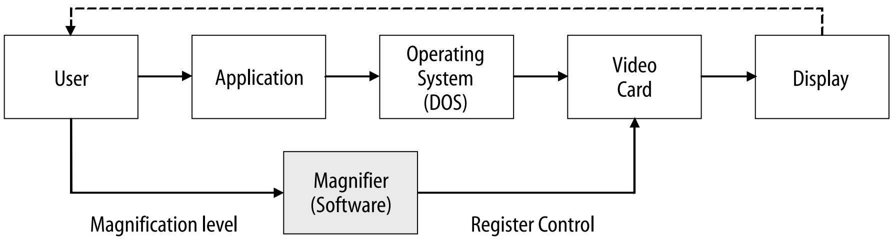


Figure 13.21. Software magnifier for DOS-based system. This controls the video card of the system independently of the operating system

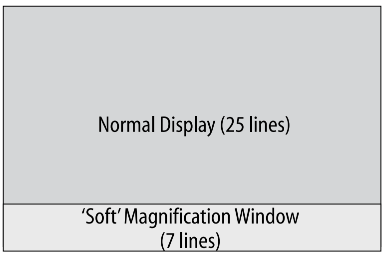


Figure 13.22. Display of the Lyons Large Print Program. This had a standard (non-enlarged) 25 line display. The bottom 7 lines of the screen were used to construct enlarged versions of the characters at the current focus. A single line of 16 enlarged characters was presented

and, by suitable manipulation of the registers, the 40 × 25 character display would scroll as the user moved the cursor, thus implementing a full-screen magnifier. In a similar way, Dolphin's Odeon system exploited the fact that PCs could have both a monochrome and a colour video card installed at the same time. The user ran standard applications in text mode on the 80 × 25 character monochrome card whilst the magnification software used the colour card to draw an enlarged image that was presented to the user on the colour card. Odeon could be operated with two monitors (colour and monochrome) connected or just with a single colour monitor (see Figure 13.23).

Later developments allowed full-screen magnifiers to be produced without the need for multiple video cards. Systems such as Dolphin's Lunar used the multiple pages of video memory provided in several of the standard video modes to achieve full screen version of their earlier HAL16 and EGA systems (see Figure 13.24).

The application and the operating system created the original image in one memory-mapped video page. The magnifier manipulated the registers of the video card so that the monitor displayed a video image from another page (the display page—which presented the enlarged image). The magnifier selected a portion of the original image and copied this to the display page. With the video card working in text (rather than graphics mode) the magnifier constructed the enlarged

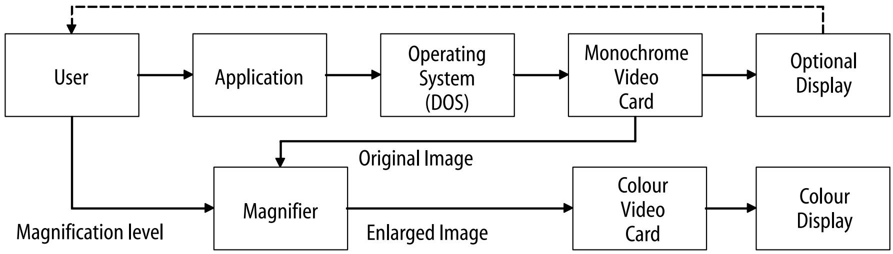


Figure 13.23. Simplified architecture of Dolphin's Odeon. The monochrome video card holds the original image; the magnifier selects a portion of this and creates an enlarged version in the colour video card. The monochrome display is optional and presented the original image

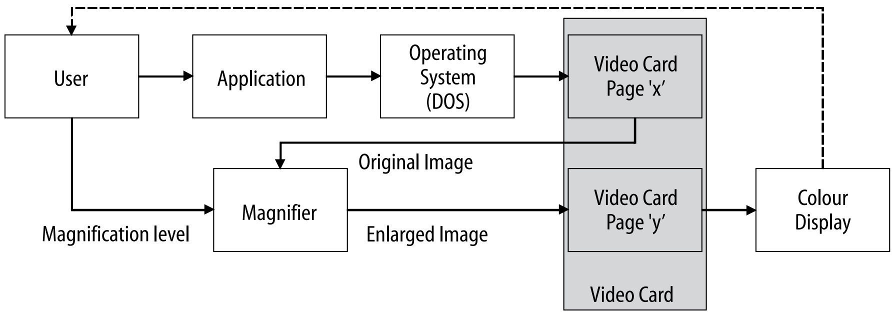


Figure 13.24. Simplified architecture of Dolphin's Luna. The original image is created in one page of the video card's memory (shown as Page 'x'), a portion is selected and enlarged and sent as the enlarged image to another page (shown as Page 'y')

image from the 256 ANSI characters available to the card. This required that for each character in the character set, the designers had to construct an equivalent enlarged character for each possible enlargement factor. Whilst a reasonable set of enlarged characters could be shown, some looked a little unusual. This problem was addressed when later video cards allowed a custom character set to be loaded. A custom character set was used that provided a set of basic building blocks from which the enlarged larger characters were constructed. This resulted in aesthetically smoother and somewhat easier to read characters. We believe that a similar approach was also used in one of the other popular magnifiers from North America, ZoomText (for DOS).

13.4.4.1.2 Magnifiers for GUI-based Operating Systems

When GUI-based operating systems such as Microsoft Windows arrived, two major changes were required to the design of magnifiers:

1. There was a requirement to magnify both graphics and text.
2. The standard video modes used changed significantly from predominantly text-based to primarily graphically-based modes.

The first point requires that the screen be magnified on a pixel-by-pixel rather than on a character-by-character basis. The second point is best considered with reference to Figure 13.18.

In Windows, an application 'writes to the screen' by making calls to the Graphics Display Interface (GDI). The GDI is an Application Programmer's Interface (API) that provides procedures to, for example, write text or bitmaps to specified locations on the screen. The GDI subsequently calls a device driver (DD), which is provided by the manufacturer of the video card. To allow the GDI to interface with a wide variety of video cards, DDs have a standardised interface and also a means of informing the GDI of the capabilities of the device so that the GDI can make the most efficient use of the card—for example, to use hardware acceleration for certain operations.

The magnifier must gain access to the original image somewhere between the application calling the GDI and the original image being created on the video card. There are four basic strategies for magnification (although most are closely related), we present these roughly in the order of the Windows graphics display chain:

- Operating system-based magnification (GDI/DD hooking).
- Device driver replacement.
- Device driver chaining.
- Video card overlay.

Operating-system-based Magnification (GDI/DD Hooking) To gain access to the original image, one approach is to intercept the calls made to either the GDI or DD and then use the information taken from the calls to construct the original image. The enlarged image is created by using the interface provided by the GDI or

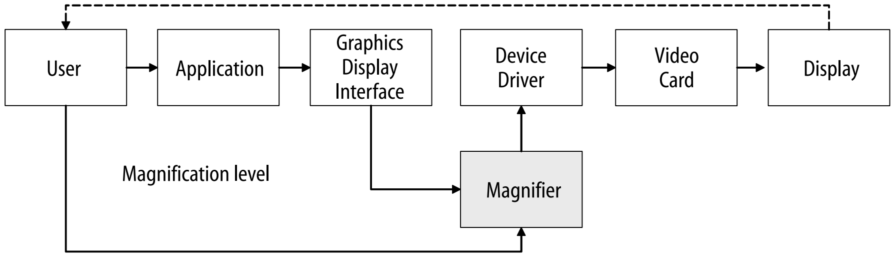


Figure 13.25. Magnifier using DD hooking. The magnifier hooks calls to the DD, builds an image consistent with these calls, and selects the magnified portion to display. It then makes a set of calls to the DD to create the enlarged image that results

DD. The process of intercepting the calls made on either the GDI or DD is referred to as hooking. A "hook is a mechanism by which a function can intercept events (messages, mouse actions, keystrokes) before they reach an application" (Marsh 1993). When calls made on the DD are intercepted this is referred to as 'hooking low'. There are several techniques that can be used to intercept GDI/DD calls, many of which are not recommended by Microsoft (but which are extensively used). One supported mechanism under Windows 95/98/Me is Microsoft Active Accessibility (MSAA) (Microsoft 1999), which provides a DDI (device driver interception) hook that allows a program to intercept the calls made on the DD by the GDI. The information contained in these calls is used by the magnifier to construct an enlarged image (as if it were the video card), which it then sends to the screen by making calls on the DD. Care must be taken because the DD is being used twice, once by the GDI whose calls are intercepted by the hook and once by the magnifier. There must be information embedded in the magnifier's calls to distinguish these from the calls made by the GDI. If not, the magnifier's calls are hooked and passed to the magnifier for subsequent enlargement—a rather unfortunate case of recursion. Hooking low is shown in Figure 13.25 and a fuller discussion can be found in (Baude *et al.* 2001).

When calls made by the application to the GDI are intercepted, this is referred to as 'hooking high'. Here the magnifier redirects the calls made on the GDI to create a screen bitmap. This means that the magnifier uses the functionality of the GDI to create and maintain the image to be enlarged. The magnified portion is sent to the screen by making calls on the DD (see Figure 13.26). Other techniques to present the enlarged image to the users are possible, *e.g.* the use of DirectX overlays to provide a surface "in front" of and obscuring the visible display, these are discussed in the section on video card overlay.

Device driver replacement In this approach a custom DD is developed that exploits the characteristics of the video card by manipulating its registers in a similar way to the approaches used for the DOS systems described earlier. For example, Sensory Software Ltd's Magnus (a Windows 3.1 magnifier from the early 1990s) replaced the standard VGA driver with a custom driver (see Figure 13.27). Note that Sensory Software produced a later magnifier that also used the name Magnus (current in 2004). The later version of Magnus uses device driver chaining. The mnemonic

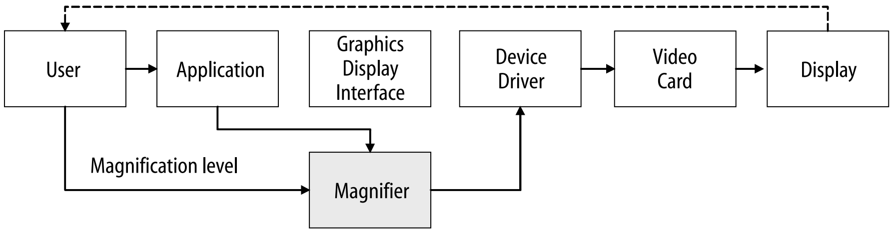


Figure 13.26. Magnifier using GDI hooking. GDI commands are interpreted and used to create a bitmap corresponding to the original image. The enlarged image is created from this. The enlarged image is usually displayed by making calls on the device driver

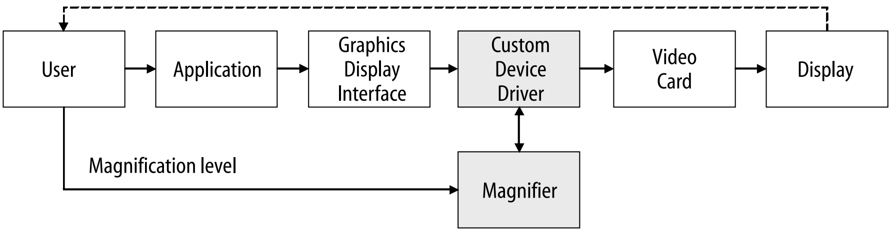


Figure 13.27. Developing a custom device driver for early Windows machines. This was a popular method of magnification for early Windows machines

VGA stands for Video Graphics Array (or sometimes Video Graphics Adaptor). VGA graphics cards were widely used in the 1990s and most of today's video cards will support VGA mode. VGA offered a number of colours from 2 to 256 colour and screen resolutions up to 640 × 480.

Magnus changed the video registers to point the memory-mapped display to a 'vacant' area of memory-mapped screen memory on the video card. This meant that the output image was generated from the 'vacant' area. The calls made by the GDI constructed the image in 'normal' display memory. The enlarged image was constructed by copying the 'appropriate' portion of the 'normal' display memory to the 'vacant' area. Magnification in the vertical direction was achieved by manipulating the VGA video registers so that lines were replicated. Magnification in the horizontal direction was achieved by copying pixels.

The approach worked well, but did have significant limitations. It replaced the generic VGA driver and therefore worked only in VGA mode. Moreover, the more advanced features of graphics cards could not be used. Both of these issues became significant drawbacks as graphics cards developed higher resolutions and more advanced features, and, most especially, when applications became incompatible with VGA.

Device driver chaining One of the features of modern versions of Windows is that a system can have multiple device drivers. These are arranged in a chain, so that the output of one device driver is routed to the input of the next. Magnification can be achieved if a custom device driver (which implements the functionality of the magnifier) is inserted between the GDI and the device driver for the card (see

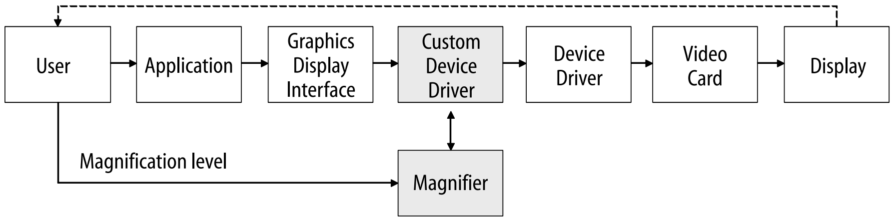


Figure 13.28. Device driver chaining. A custom device driver is inserted between the GDI and the device driver

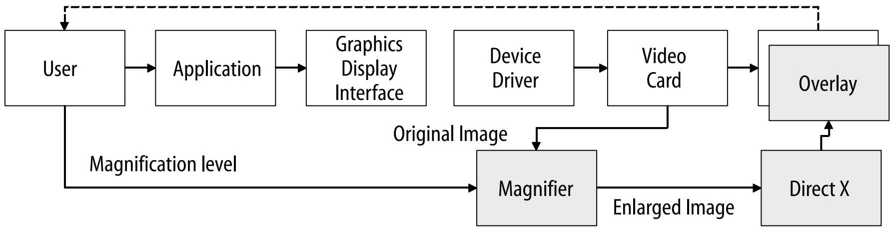


Figure 13.29. Overlaying the enlarged image. The original image is taken from the video card which is created by Windows in the normal way. Direct X is used to overlay the enlarged image over the displayed original

Figure 13.28). In many respects this is similar to hooking low (at the DD level) in that the magnifier must interpret calls made to the DD to create the original image, select an area and magnify it, and pass the enlarged image to the standard DD by making calls on it. The advantage of this approach is that hooking low is not supported by later versions of Windows (XP and 2000) and DD chaining is.

Video card overlay Rather than the magnifier intercepting calls made in the graphics display chain of Windows, the magnifier can allow the Windows operating system create the original image on the video card in the normal way. The enlarged image is overlaid on top of that created by the video card, obscuring the original image. This simplifies the design of the magnifier because the magnifier does not have to create the bitmap; the original image can be taken directly from the video card. Microsoft Direct X (Direct Draw) overlays provide a means by which one image can be laid over another obscuring the image that lies beneath. In this approach, the magnifier accesses the video card to obtain the original image and displays the enlarged image 'on top of' the original image by using a Direct X overlay that fills the screen obscuring the original image. This approach is illustrated in Figure 13.29 and is discussed in more detail later when we present the architecture of a modern screen magnifier.

13.4.4.2 Discussion of Methods

We believe that the early methods that involved custom hardware are no longer practical because custom hardware is expensive compared with software-only solutions. Similarly, the methods that directly accessed the registers of the video card are generally no longer valid given the wide variety of video cards that are now available. This also militates against the use of custom device drivers because a designer either has to develop a custom DD for each video card on the market or has to develop a reduced functionality generic driver that has lowest common denominator functionality (for example, a generic VGA driver). The former involves significant development and maintenance costs and the latter may significantly reduce the range of applications that can be supported and the performance of the system. This leaves three methods that are, or have been recently used, in commercial magnifiers.

Hooking has been widely used. However, 'hooking low' (DD hooking) is not supported in Windows 2000 and Windows XP and thus can only be used for the now obsolete Windows 98 family. We believe that 'hooking high' (GDI hooking) has been widely used and is still used by some commercial magnifiers. This is an attractive option, but does have some problems. There may be some noticeable reduction in the performance of a computer system running such a magnifier because interpreting GDI commands in the screen magnifier is computationally intensive and because memory must be allocated to build the original image, making demands of system resources. Moreover, GDI hooking is unsupported and documentation is sparse, making implementation difficult. Nonetheless, as noted earlier, this approach is widely used.

The other approaches are also used in commercial magnifiers and in the next section we discuss a magnifier based on the video card overlay technique, which was used in the commercial magnifier Magnice (Sensory Software Ltd). This approach suffers from one disadvantage in that is requires that the user's video card supports overlays. Not all cards do so, and whilst the number of cards that do is increasing, there can be problems with legacy machines and cards that restrict the way in which overlays are used (for example, restricting the resolution of the overlay). Nevertheless, it remains a valid approach to magnification.

Finally, device driver chaining is commercially used and is a relatively low overhead approach.

13.4.4.2.1 The Architecture and Implementation of a Modern Screen Magnifier

This section presents the architecture of Sensory Software Ltd's Magnice screen magnifier, which used the video card overlay technique. We choose to present the architecture of this magnifier (rather than its successor Magnus, which uses device driver chaining) because further details are available to the interested reader (Blenkhorn *et al.* 2002). The method of accessing the original image and displaying the enlarged image differs between Magnice and Magnus but other considerations remain the same. The architecture of Magnice is shown in Figure 13.30.

The basic operation of this screen magnifier is that the operating system and applications construct the original image on the video card. For most applications and the operating system, this is through the GDI/DD chain discussed earlier. However, this may be bypassed when Direct X's Direct Draw is used and one of the advantages of the approach discussed here is that even if an image is constructed that bypasses the GDI/DD chain it is often magnified correctly. The magnifier

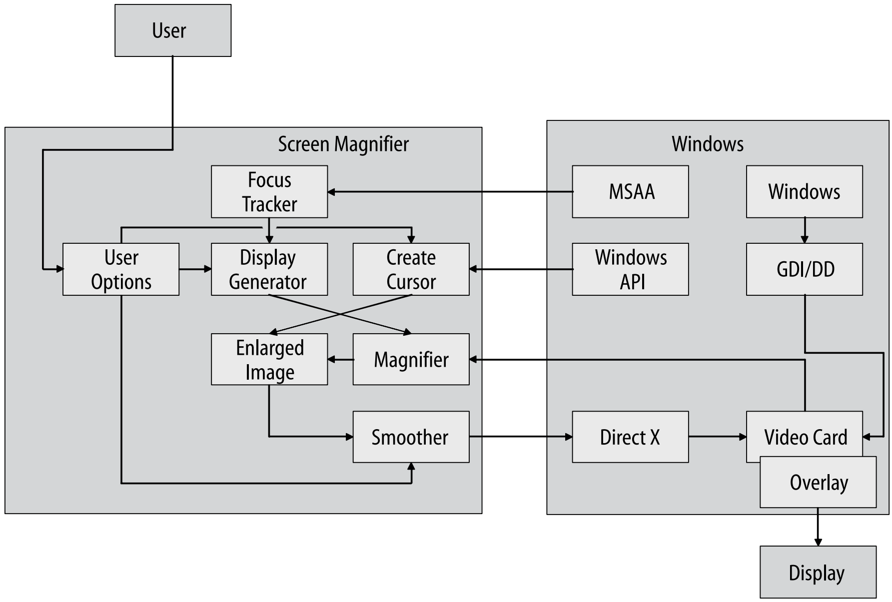


Figure 13.30. Simplified architecture of a Magnice

accesses the original image and magnifies it to create the Enlarged Image, applying the settings supplied from the user (*via* User Options to Display Generator). The Enlarged Image is smoothed (if required, this is a user setting supplied to the Smoother from User Options) and Direct X's Direct Draw is called to display the Enlarged Image as an overlay that obscures the original image when displayed.

There are a number of issues that have not been discussed in our consideration of magnifiers this far. We present these below; they are:

- How does the magnifier track the user focus?
- How does the magnifier enlarge the mouse pointer (the cursor)?
- How is magnification actually achieved?

Tracking the focus As noted earlier, the magnifier needs to track the focus irrespective of whether focus changes due to user, application or operating system actions. The Focus Tracker uses two means to track the focus. It can use a mouse hook to determine when the mouse has moved. A mouse hook simply intercepts events that correspond to changes in portion and status of the mouse buttons. Microsoft Active Accessibility (MSAA) is used to determine other focus changes. MSAA "is a suite of technologies for the software developer that lets applications and objects communicate and automate more efficiently. MSAA allows applications to cooperate effectively with accessibility aids, providing information about what they are doing and the contents of the screen" (Microsoft 1999). MSAA signals when significant events occur, and the Focus Tracker listens to these and determines when a focus change of interest has occurred. Thus, the Focus Tracker is aware of when a window pops up and takes focus; it is also aware of which element on the window has the focus. The Focus Tracker informs the Display Generator of the current area to be magnified. The Display Generator uses this information and that gathered from the User Options to control the Magnifier.

Mouse pointer enlargement In many computer systems, the display of the mouse pointer is carried out directly by the hardware graphics card and normal methods of graphics display are bypassed. In many respects the standard method of the displaying of the mouse pointer is similar to the use of overlays used for magnification in Magnice. The mouse pointer is overlaid on top of the image, but, unlike the magnification overlay, it does not overlay the whole original image, only the part that corresponds to the position of pointer. When Magnice captures the original image, the mouse pointer is not present on that image. This means that the magnifier has to generate its own mouse pointer and place it on the enlarged image at the correct position and with the correct degree of enlargement. Note that this discussion does not only apply to Magnice taking the original image directly from the video card, it also applies to methods that use hooking and device driver chaining.

To construct an enlarged mouse pointer the following steps are required:

1. The mouse pointer needs to be located. This can be done by the standard Windows Applications Programmer Interface (API) call (GetCursorPos) which returns the position of the cursor.
2. The bitmap corresponding to the mouse pointer needs to be determined. This will change dependent upon the application or operating system's state. For example, the standard arrow cursor will be changed to an hourglass when an application or the operating system is busy. Moreover, the user may have chosen to use an alternative, non-standard set of cursors, and these, rather than the standard set, need to be enlarged. Access to the bitmap is achieved by calling a standard Windows API function (GetIconInfo).
3. A magnified version of the mouse pointer now needs to be created. This can be achieved by copying a magnified version of the mouse pointer's bitmap (with appropriate transparency) onto the magnified area of the screen.
4. The system needs to prevent the graphics card from displaying the hardware-generated unmagnified mouse pointer. This can be achieved by making a further API call at regular intervals, (*e.g.* when the mouse moves), that hides the mouse pointer (SetCursor(NULL)). This needs to be done regularly, because every time Windows changes the mouse pointer for an application, *e.g.* when a pointer turns to an hourglass, the new cursor is displayed.

With careful placement of the magnifier-generated mouse pointer (so that it matches the position of the system's mouse pointer), the enlarged mouse pointer will appear and, from the user's perspective, appear to work as normal.

Magnification We have spent a considerable amount of time discussing how a magnifier gains access to the original image and how it presents the enlarged image,

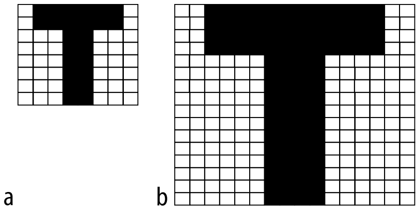


Figure 13.31a,b. Magnification is copying rows and duplicating pixels. The 8×8-pixel image **a** is enlarged by a factor of 2 by duplicating every row in the vertical direction and every pixel in the horizontal direction to become the 16×16 image **b**

but we have not yet discussed how the enlarged image is created from the original image. The reason for this is that, in many respects, this is the easy part!

Magnification is achieved by duplicating pixels. In the vertical direction, rows are duplicated by the enlargement factor. In the horizontal direction, pixels are duplicated by the enlargement factor. So, if the enlargement factor is $2\times$, each row is duplicated in the vertical direction and each pixel in the horizontal direction (see Figure 13.31) for an example). Magnification is speeded up by using fast pixel and area copying techniques ('blitting').

The only difficult issues to deal with are non-integer enlargement factors, for example an enlargement factor of 1.5, for it should be noted that the vast majority of magnifiers only provide integer enlargement factors. The authors are not aware of how this issue is handled by commercial implementations. However, it is likely that a large version of the enlarged image is created using standard row and pixel duplication techniques and that this is sub-sampled when the enlarged image is presented on the screen.

13.5 Screen Readers

13.5.1 Overview

The purpose of a screen reader is to give a blind person access to, and control over, a computer system. The information that is normally relayed on the screen will be presented to a blind person either through synthesised speech or through a Braille line (on a refreshable Braille display). For the purposes of this chapter we will assume that we are dealing with a talking screen reader (*i.e.* one that uses speech as output); we will briefly consider the use of a Braille line later. The reason for doing so is that the vast majority of screen reader users use speech because of the relatively high cost of a Braille line and the relatively low levels of Braille literacy (especially amongst those who become blind later in life). This is not to say that Braille displays are not of considerable benefit to those who can afford and can use them.

First it is important to understand how a blind person will interact with a computer system using a screen reader. The user will use the keyboard to control the computer and (even on Windows-based computer systems) will not use a mouse. This is because the use of a mouse involves a closed visual loop where the user sees both the position of the mouse pointer and where they wish to move the focus. This closed loop is does not work if the mouse pointer cannot be seen. The implication of this is that the user must learn how to control their computer solely through the keyboard. GUI-based operating systems, such as Windows, may appear to be wholly dependent on the mouse. However, they can be controlled through the keyboard using a large number of keyboard shortcuts.

Like a screen magnifier, a screen reader must track the focus. The information at the current focus is spoken to the user. The screen reader will not just speak the text associated with the focus but will also give some indication of the type of the element that has the focus. For example, when a menu item has the focus, the name of the menu item will be spoken (for example the 'New' operation in Microsoft Word) followed by the text 'menu item'. Similarly, when a checkbox is reached the name of the checkbox is spoken, followed by its status (checked or unchecked) and then the fact that it is a checkbox. So in Figure 13.32, the screen reader will read "Novice Mode, unchecked, checkbox" when the focus is on the left hand checkbox and "Use Music, checked, checkbox" when focus is on the right hand checkbox. Note that the speech given here is for Sensory Software Ltd's LookOUT screen reader. Other screen readers will produce similar speech, but may not necessarily give the same information in the same order.

The best way of gaining an understanding of the way in which a screen reader works is to experience one in use. It is suggested that you follow Project P.2 at this point. Even if you do not have access to a Windows machine, there is some explanation in this project which should be read to assist your understanding of what screen reading involves.

The primary goal of a screen reader is to present to the user information at the current focus. The way in which the information is spoken is dependent on the type of information at the focus and the user's current task. We will explore this issue further in the sections on usability issues later. One important point to note is that a screen reader should try to give the maximum amount of information with the minimum amount of speech. It is a general observation that the more speech that is given, the more frustrating a user finds the system to use. So, the real goal is to give only as much information as is strictly necessary for the user to accomplish their task.

Figure 13.32. Checkboxes. The screen reader will read "Novice Mode, unchecked, checkbox" when focus is on *checkbox on the left*. It will read "Use Music, checked, checkbox", when the focus is on the *right hand checkbox*

13.5.2 The Architecture and Implementation of a Screen Reader

In this section, we present a brief overview of the history of screen readers to give the reader a greater understanding of the technical issues involved in the implementation of a screen reader. We then present a simplified view of the architecture of a modern Windows screen reader (Sensory Software Ltd's LookOUT).

13.5.2.1 Historical Approaches

Blind people have used computer systems since the late 1950s. In the days of punch card input, some blind users could read the cards by touch. Output was generally through Braille, although later primitive speech synthesisers were used (see for example to work of Sterling *et al.* 1964 and Rubenstien and Feldman 1972). Later systems introduced scrolling Teletype input that could be accessed using talking terminals. Although these systems gave access to computer systems they were not screen readers in the current sense of the meaning. Screen readers only really came into being with the introduction of character user interface (CUI) systems where the information on the screen was memory mapped, *i.e.* there was an area in memory in which the characters and their attributes (*e.g.* bold, inverse, colour, *etc.*) were stored. Screen readers were developed for the BBC Micro, NEC 8201 'laptop', and for CP/M. However, significant numbers of blind users really started to use personal computers with the introduction of the IBM PC (and compatibles) and MS-DOS (or just DOS). The techniques that we describe below for DOS are similar to the ones for the other CUI systems. Although there were different "resolutions" possible for the rows and columns of text on these systems, we restrict the description to the most common on used, namely, 80 columns of 25 rows.

In common with modern screen readers, these screen readers had to accomplish three basic tasks:

1. To be able to determine the identity of a character on the screen.
2. To be able to echo the keyboard (*i.e.* to read back to the user what they have typed) and filter out keystrokes that correspond to commands to control the screen reader.
3. To be able to determine the current focus. In early systems this corresponded to the cursor position.

13.5.2.1.1 Identifying a Character

In a DOS system, the computer would use two bytes, one to store the character and the other to store the attribute. The characters in each column and row were stored in sequential locations. To determine which character was at a particular position on the screen, the screen reader had only to "peek" (*i.e.* read) at the memory locations corresponding to the current focus (given by cursor position).

13.5.2.1.2 Keyboard Issues

Keyboard input for DOS program was *via* two low level operating system calls (at the BIOS level). One would return the character and remove it from the input buffer; the other would return the value of the character without removing the character from the input buffer. When a key was hit, the screen reader needed to determine the identity of a character that was input. The screen reader could then take three actions dependent on the character or the current sequence of characters. First, the character could be passed to the application and its value read out. For example, typing an A cause 'ay' to be spoken—this is called keyboard echo. This is appropriate when, for example, the user is typing text into a word processor. Second, the character might correspond to movement of the cursor, for example, a 'cursor up' command. In this case, the identity of the character would again be passed to the application, but the screen reader would not echo the key but speak in a manner appropriate to the action. So, for example, 'cursoring up' in a word processor would cause a full line of text to be read. The conventions for reading text following a cursor operation that were established in the days of DOS systems are still followed by the vast majority of screen readers and self-voicing applications today—see Project P.2. Finally, some characters would correspond to commands directed at the screen reader, for example, to read the current word or to change some characteristic of the speech synthesiser. In this case, the character is 'swallowed' by the screen reader and not passed to the application.

By redirecting the system's keyboard to the screen reader, keystrokes could be filtered and appropriate actions could be taken. One slight complicating factor was that some programs only examined the status of the text in the input buffer and rather than calling the low level operating system functions to remove it. Such systems needed to access the keyboard buffer directly to remove characters. Consequently, appropriate heuristics needed to be applied to both filter (hook) input and examine the keyboard buffer.

13.5.2.1.3 Tracking the Focus

Initially, DOS programs would indicate the focus by use of the hardware cursor; its position was readily accessed by a call to the low-level operating system (the BIOS). If the cursor were moved by using one of the cursor keys, a suitable response was given, *e.g.* cursor left and right would read the next character, control-cursor left and right would read the next word, and up and down arrow would read the line (in the same way as described in the project P.2 given at the end of this chapter). Later some systems abandoned the use of the hardware cursor and changed the attribute of the character to invert its value (*e.g.*, a white character on a black background would be changed to a white character on a black background to signal the position of the cursor). In this case, the whole of the screen's memory map had to be searched for the inverted attribute to determine the cursor position (as noted above, the screen's memory used one byte to hold the character and another to hold its attributes). At the same time software would increasingly use a highlighted bar to indicate menu choices (previously these operations would be by keyboard commands). Here a similar technique to that used to track the 'inverted cursor' was used.

13.5.2.2 The Architecture of a Modern Screen Reader

This section gives an overview of the architecture of a modern screen reader (Sensory Software Ltd's LookOUT). Further details of this architecture have been provided by Blenkhorn and Evans (2001) and Evans and Blenkhorn (2003).

The architecture of the LookOUT screen reader is shown in Figure 13.33. Before we look in detail at its operation, we need to identify some of the characteristics of its major components. We will briefly consider hooking, MSAA, OLE and the off-screen model.

13.5.2.2.1 Hooking

There are a number of circumstances when a screen reader needs to intercept input before it reaches an application and output before it reaches the screen. For example, as we have seen earlier, screen readers need to intercept keyboard input to obtain the commands that control the screen reader and prevent these from being routed to the application. Screen readers also need to intercept the calls made by applications and the operating system on the operating system's graphics system to be able to determine the information that is being written to the screen.

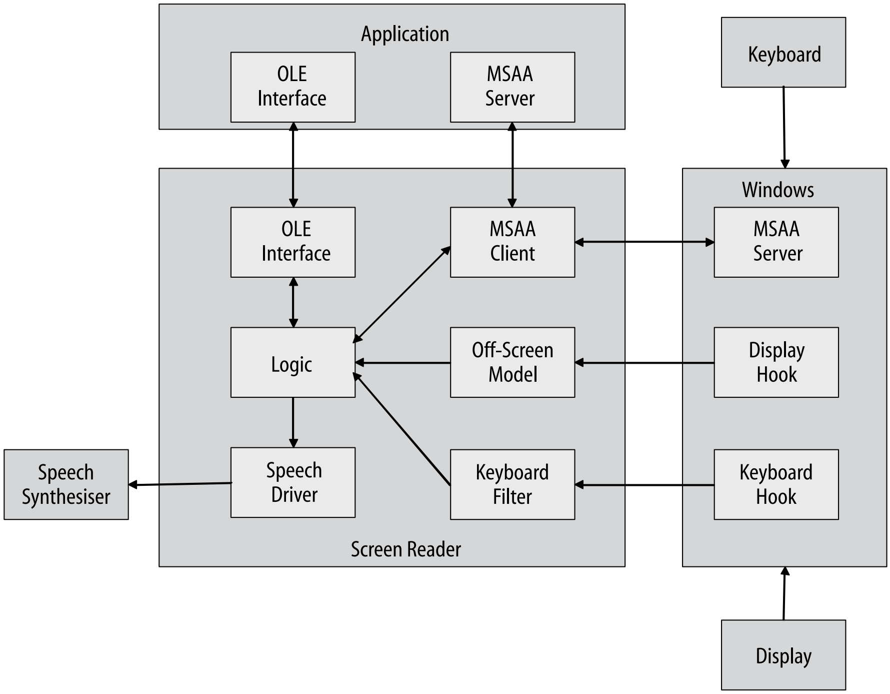


Figure 13.33. Simplified architecture of LookOUT

The interception of information in this way is achieved by using a hook. We have met hooks earlier in this chapter when we considered GDI and DD hooking in a magnifier. A hook is simply a method by which data being passed to and within the operating system can be extracted and used by some external application. Data that is obtained through a hook can be filtered. This means that some data is extracted and used by the screen reader and other data is passed through to the applications and the operating system. So, for example, keyboard commands to the screen reader are filtered out.

13.5.2.2.2 Microsoft Active Accessibility (MSAA)

As noted earlier (when we discussed the architecture of a modern screen magnifier), MSAA provides a means by which applications and the operating system can generate events and expose their contents to other applications through a standard interface. It is the application's responsibility to provide this interface—an MSAA server. A screen reader (or, indeed, any other assistive application) can take advantage of this interface (by implementing an MSAA client) to determine what information an application is presenting to the user. Not all applications provide an MSAA server (for example, the client areas of Microsoft Word and Excel are not supported). However, the screen reader can generally get some useful information if the application uses standard interface components such as pull down menus and controls. This is because these are implemented by the operating system and because Windows itself provides an MSAA server from which this information can be extracted.

13.5.2.2.3 Object Linking and Embedding (OLE)

OLE is an automation technology developed by Microsoft that allows applications to work together. An application that has an OLE interface exposes its Object Model. The Object Model can be used by a screen reader to extract data from an application and also to control the behaviour of the application. This is very useful for applications that do not fully support MSAA. For instance, the area in which a user types in Microsoft Word is not exposed through MSAA, but Word has an OLE interface through which a screen reader may obtain data. OLE also allows the screen reader to control the application and there are some circumstances where this is appropriate. For example, an application may not have a good set of keyboard shortcuts. These can be provided by a screen reader, which can hook and filter keystrokes, to determine which keystrokes are 'simulated shortcuts' for the application and control the application directly through its OLE interface.

In order for a screen reader to work with OLE, the screen reader must be aware of which application it is communicating with and the characteristics of the OLE interface provided by the application. This means that the screen reader must have application-specific code for each application that it drives through OLE. The code to interact with common applications is hard-coded into most screen readers. Control of other applications through their OLE interfaces may be achieved by writing additional code for those screen readers that use scripts (see the section on Scripting later).

13.5.2.2.4 Off-screen Model (OSM)

Screen readers need to be able to determine the information that is displayed on all parts of the screen. MSAA and OLE can satisfy many of the demands for information that a screen reader may make. However, MSAA and OLE cannot satisfy the requirement that a screen reader will interact with almost all applications under all circumstances. Thus, a screen reader needs to maintain its own version of the information that is displayed on the screen. This is achieved by the screen reader hooking the calls that applications and the operating system make on the graphics display system. This can be achieved by GDI hooking; this process was described screen magnifiers earlier in the chapter. By intercepting the calls, the screen reader can build an off-screen model (OSM). This is data structure that has an entry for every pixel on the screen. Each entry holds the attributes of the pixel. The attributes do not just contain the colour value of the pixel, but indicate the larger structures to which a pixel belongs. So for a pixel that is part of a character, its attributes would include the value of the character, its font, it display characteristics (bold, italic, underlined, *etc.*), and the string that contains the character. If the pixel were part of a screen element, for example, a button on a form, this information would also be included. By holding information of this type for every pixel on the screen, the OSM presents an interface that allows the screen reader to determine the attributes of a pixel. For example, a screen reader can query the OSM to determine whether a pixel is part of a character or not. If it is, the screen reader can obtain information about character's value, its font and its dimensions. The OSM will also be capable of informing the screen reader that significant events have taken place (in the same way as MSAA). For example, the OSM can send an event to the screen reader whenever the 'blinking caret' in an application moves (*e.g.* the text entry point in Word). Some screen readers increase the blink rate of the 'flashing caret', probably in an attempt to make the detection of caret movement easier through an OSM.

13.5.2.2.5 Operation

The user controls LookOUT through the keyboard. With regard to keystrokes, LookOUT operates in the same way as the DOS screen reader described earlier. They need to be interpreted as commands or passed onto the application or operating system. LookOUT uses a Keyboard Hook and Keyboard Filter to extract the keystroke combinations that form the commands. The other keystrokes are passed to the operating system, which places them in the Windows keyboard buffer.

LookOUT uses two major sources of information, Microsoft Active Accessibility (MSAA) and its internal off-screen model (OSM). LookOUT always uses MSAA as its first source of information. MSAA is used as a source of events that indicate, for example, that a pop-up window has appeared. MSAA is also queried by LookOUT so that LookOUT can determine the type of information. For example, MSAA can be used to determine the type of element on a form, *i.e.* whether it is a checkbox, radio button, *etc.* MSAA events are also handled by LookOUT in order to provide context sensitivity.

As noted earlier, not all applications support MSAA and indeed a small set of the standard Windows controls are not supported by MSAA. When information cannot be obtained from MSAA, LookOUT uses the OSM. The OSM obtains its information by hooking the calls made to Windows Graphics Display Interface (GDI). By interpreting these calls, the OSM can establish the type of information held at every pixel on the screen. The off-screen model used by LookOUT was developed jointly by Microsoft, ONCE, Baum Products GmBh and euroBraille SA and this OSM can be obtained as part of an open source project (SOS, undated).

In some circumstances, it is difficult for the screen reader to determine information from the OSM. The classic example is determining the selected cell in Microsoft Excel. This is indicated visually using quite subtle means (providing a highlighted box around the cell and setting the row and column names to bold—see Figure 13.1). The screen reader must be able to interpret these rather subtle visual cues and doing so can be quite computationally intensive. However, if an application provides an OLE interface and the screen reader is aware of which application it is dealing with, the information can easily be obtained by making a call on the object model. Continuing with the Excel example, it is relatively simple to use Excel's OLE interface to find which cell is currently highlighted. This technique is computationally efficient and reliable.

Given that the screen reader has access to information about what is being displayed by the application, the Logic part of the screen reader (see Figure 13.33) determines what should be spoken. This is then sent to as a speech string to a text-to-speech synthesizer. LookOUT will drive any speech synthesizer that is installed on the system that complies with the Microsoft Speech API (SAPI 4 or SAPI 5). Determining what to speak depends upon the current focus, the user's actions and preferences and, in certain circumstances, the application that is being driven. Complex heuristics are used to determine what to speak and when.

13.5.3 Using a Braille Display

As noted earlier, a Braille Display (or Braille Line) can be used instead of speech output. A Braille display consists of a line of refreshable Braille cells (between 20 and 80)—see Figure 13.34. Standard Braille cells have six dots to present a character, but most Braille displays have eight dots. The additional dots are used to indicate attributes of the characters (for example, whether they are capitalised or not) and to present a cursor. The cells are refreshable in the sense that they are made up of an array of retractable pins. The set of pins that is required to represent a character are elevated to represent the character; the others are retracted. The Braille display typically presents a line of text, with the focus presented somewhere on the line and indicated by a flashing cursor (the bottom right dot of the cell rising and falling).

There are some significant advantages to using a Braille line. The user can easily review information—something which is much less easy and convenient in a speech-only system. The user also gets a more accurate representation of text on

Figure 13.34. Refreshable Braille display of 40 cells (image courtesy of BAUM Products; www.baum.de)

the screen. This can be very useful when proofreading. When using speech it is not possible to easily detect the difference between the homophonic heterographs (red (colour) and read (past tense of the verb 'to read')). The disadvantage of a Braille line is the relatively high cost (generally considerably more than a highly specified personal computer). Also, as noted earlier, only a relatively small proportion of the blind population (around 5%) can actually read Braille.

13.5.4 User Interface Issues

In this section we introduce two additional user interface issues. The first is concerned with the specialisation of screen reader behaviour for different applications by the use of scripts. The second is concerned with the information that the screen reader speaks to the user.

13.5.4.1 Scripting

As noted earlier, the screen reader logic determines what to speak dependent on focus, user application and the application. Some screen readers, including LookOUT and Freedom Scientific's JAWS (FrSc, undated), allow the screen reader to be specialised for applications by writing scripts. Scripts are programs written in a programming language (LookOUT uses Visual Basic Script and Jaws uses its own proprietary scripting language) that override a screen reader's default behaviour. In effect, this means that the scripts determine which information is sent to the speech synthesiser under certain circumstances. Scripts are developed for a particular application, will be loaded whenever that application is run and will be applied whenever the application has the focus. Scripts have access to the same sources of information as the screen reader (*i.e.* information from MSAA and the OSM), which they can obtain from the screen reader. In addition they may gain access to other sources of information, for example by accessing an application through its OLE interface. Generally, scripts gain access to information and filter it before it reaches the screen reader. So, for example, a script can filter keystrokes before they are passed to the screen reader. This can be useful if an alternative keyboard interface to an application is required. For example, we have found it useful to control Windows media player from the numeric keypad, with the media player commands being mapped onto numbers. This is achieved by a script that intercepts the numeric keypad characters and applies an action (*e.g.* play, stop, pause, skip, *etc.*) to the Windows Media player from the script.

Scripting can be very useful for specialising a screen reader to interact with a particular application. Scripted screen readers make very extensive use of this technique. It is possible for end-users or their colleagues to write scripts for screen readers to ensure that the screen reader works well with a particular application. However, it has to be said that writing screen reader scripts is a relatively specialised and complex task and that few people write scripts.

Many screen readers are not scripted. If a screen reader is not scripted then little can be done to specialise the screen reader for a particular application. It is our experience that screen readers that are not scripted tend to work well with a wide range of common applications and may also work well with more esoteric applications. Indeed screen readers that are not scripted often work better on unusual applications than scripted screen readers that have not had a specialised script written. However, with a scripted screen reader there is always the opportunity to write a script and tune its performance.

13.5.4.2 Information Production

It is obvious that a screen reader should speak information that is useful to the user. However, there is a question as to what information is useful. To help to explain the issue, we will take a specific example. Suppose that a screen reader user is using Word and they decide to spell check the document. This will produce the form shown in Figure 13.35.

The screen reader will be aware that the form appeared (*via* MSAA or the OSM) and that the focus is on the first of the suggestions. It must now decide what to speak. It must obviously tell the user that the Spelling and Grammar Check Window has appeared by saying something like "Spell Check Window". It must then decide what to speak next. One strategy is to acknowledge that a new window

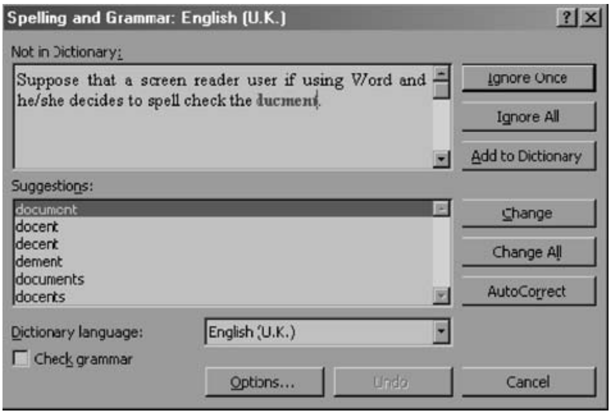


Figure 13.35. Spell Check window in Word

has appeared and to describe it to the user. So, for example, a screen reader could inform the user that the form has nine buttons, two text boxes, a check box and a drop down selection box. It could also give their identities before stating which element was selected. This approach, *i.e.* to describe a new window or form when it appears, is used by a number of commercial screen readers. The reasoning behind this decision seems very logical; a new graphical item has appeared on the screen and the user should be informed of what it is presented. However, it does lead to a lot of speech and an experienced user is probably aware of the options available from forms that appear in applications that they use regularly. In the case of spell checking, the user will probably want to know immediately which word is misspelled and what the replacement options are. However, to find out which word is misspelled they will have to move the focus to the upper text box. An alternative approach is to recognise that when a user is spell checking, they will probably want to find out which word is misspelled and what the options are as quickly as possible. So in this approach the following text results "Spell Check Window, Not in Dictionary, *ducment*, suggested alternative document, d-o-c-u-m-e-n-t". This speech is taken from Sensory Software Ltd's LookOUT Screen Reader and other screen readers will give similar output. Clearly, this is a very task oriented description, which spends no time describing the form that has appeared or the options available to the user. The user can determine what options they have by 'tabbing' from control to control in the normal way. So, if the user wants to know more, they must make the effort to explore. This may be a problem for novice users, who will have to search to find out what options they have. However, this may be better than describing the window in that the user is immediately informed of the task of this form.

The example demonstrates that are there two contrasting views of what a screen reader is. On the one hand, a screen reader can be thought of as a tool that allows blind people to get a detailed description of what a sighted user sees on the screen. On the other, a screen reader can be thought of as a tool that presents a blind person with an interface to a standard application, as if it were an application designed for blind people. Returning to the example, the first approach tells the blind user that 'this is a spell checker window and the sighted person can see the following things'. The second approach tells the blind user that they are using a spell checker and presents the information that is relevant to the most common action that is undertaken with this form.

In practice, screen readers tend to incorporate both approaches, although traditionally most screen readers have acted like inspection tools rather than the facilitators of talking applications. The second approach relies heavily upon detailed and complex scripting and will thus only work where scripts exist. When an application is met that is not supported by a script, the screen reader must revert to the traditional inspection tool approach. It is the authors' experience that many users, especially those who are not experts in the use of screen readers and computers, prefer the second approach. This is because the system speaks less, and when it does speak, it is generally targeted to the user's current task.

13.6 Hybrid Screen Reader Magnifiers

Some users have sufficient functional vision to read some information from the screen when it is magnified, but also wish to have the support of a screen reader at the same time. This has led to the development of hybrid screen reader magnifiers that provide the characteristics of both tools in the same package. Little needs to be said in terms of user experience and system architecture. Both combine the features of their stand-alone equivalents. The one issue to note, perhaps, is that most hybrid screen reader magnifiers derive primarily from either a screen reader or a screen magnifier and it is the operational characteristics of the tool from which the hybrid system derives that dominates its characteristics. So, a hybrid system that derives chiefly from a screen reader is likely to have better screen reading functionality than it does magnification capability and *vice versa*. A further issue to note is that the user of a hybrid system may well be a mouse user. This means that the screen reading part of the system must respond to changes in focus that are prompted by mouse actions. However, screen readers are designed to detect changes in focus, whatever the source, so this is not a major issue.

13.7 Self-magnifying Applications

There are a significant number of applications that allow a user to configure the presentation of textual information. Control can be given over many features including character size, inter-character spacing, inter-line spacing, colours and font size. The work reported by Fairweather *et al.* (2002) provides a good introduction to the range of alternatives to the presentation of text for people with visual impairments and other disabilities.

In many applications (for example, common word processors and document presentation tools), the user can zoom in and out thereby controlling the size at which images are presented. Whilst there is good support in many applications for enlarging the client area, few applications provide much control over the size of menu items and icons. This can be a limitation for some users. When designing an application for people who may have visual impairments and who will use the application without a screen magnifier, it is important not to forget to provide large text options for menus and large icons where needed. An example of a simple word processor that provides such support is the EdWord word processor. This is the subject of the project P.4 in the projects section at the end of the chapter.

13.8 Self-voicing Applications

There are a number of self-voicing applications. In many respects, the user experience of a self-voicing application is very similar to that of using a screen reader with a standard application. There is one major difference. The designers of self-voicing applications have full control over the application and the speech. This means that the designers have a full understanding of the context of all information that is sent to the user, so, if it is thought desirable, they can change the speech synthesiser from male to female to indicate a different type of information.

One application in which this type of approach has been used is in IBM's Home Page Reader (IBM, undated). Home Page Reader is a web browser that is aimed at visually impaired people and provides speech and magnification. One tactic that has been used in Home Page Reader (and indeed in other stand-alone web browsers) is to distinguish between text and links by changing the voice in which the information is spoken. By default, screen readers do not change voices to signal different types of information. However, if appropriate application scripts are developed, this is supported by some screen readers working with particular applications.

There is a question as to why a user would use a self-voicing application in preference to using a standard application and a screen reader. After all, an independent blind computer user will almost certainly have to use a screen reader to get the self-voicing application to run. There are three circumstances when self-voicing applications are appropriate. First, when the application is presenting complex visual information that can be simplified if it is presented specifically for visually impaired people through speech. A good example of this is web browsing, which many users find difficult through a screen reader. Second, where the user possesses sufficient functional vision to run applications by sight, but who can benefit from having speech augment their visual experience when using a complex application. Again, web browsing a typical application and this approach is often used with older users. Finally, some users will not be able to use computer systems independently. This is typically due to cognitive as well as visual impairments. In this case, a helper will select and start the application and the user is presented with an easy to use application. An example of this type of system is the EdWord word processor that is the subject of the project P.4 in the projects section at the end of the chapter.

The user's experience of using a self-voicing application is quite similar to that of a screen reader user accessing that application designed for sighted users. When 'cursoring around' the self-voicing application, screen reading conventions discussed earlier are almost always followed (*e.g.* cursor up, read line; cursor right, read character; shift-cursor right; read word).

From a designer's perspective, self-voicing applications are relatively simple to design. Unlike screen readers that must determine what is being written to the screen by hooking the graphics display chain, the designer of a self-voicing application knows exactly what is to be spoken; it is presented by their application. To speak, the designer simply sends a string to a speech synthesiser. There are some pitfalls to avoid when writing self-voicing applications. Novice designers often make their system far too verbose. Follow the maxims of 'minimum speech, maximum information' and 'do not speak unless absolutely necessary'. Keep messages short and do not provide instructions—leave these for a help file. Another pitfall is not interrupting speech messages when a user changes focus. The user should always be informed of the latest focus information; this may mean interrupting an earlier speech message.

13.9 Application Adaptors

Some applications are difficult to use through a screen reader. For example, reading web pages is very difficult for some screen reader users and complex applications that are heavily reliant on visual interfaces can also prove difficult to use, *e.g.* Microsoft Outlook. One way of getting around these problems is to build custom self-voicing applications. A number of stand-alone web browsers have been built for blind people. However, the developers of such systems face the problems of keeping their tools up-to-date, matching the features, protocols and standards of commercial product aimed at the mass market. This is very difficult; imagine designing, implementing and maintaining an application that had the same level of functionality of Microsoft Outlook.

So there is a potential problem. On the one hand, some applications may prove too difficult to use for some users using a screen reader. On the other, the effort involved in designing self-voicing applications that mirror the functionality of commercial product is prohibitive. One approach to this problem is to develop an application adaptor. This is a program that sits between an application and a screen reader. An application adaptor will interact directly with the standard application using OLE and it will present the information extracted from the application to the user in a 'screen reader friendly way'.

The authors have adopted this approach for two applications. We have developed an interface that sits between Microsoft Internet Explorer and a screen reader (WEB, undated). This interface (Webbie—see the project P.3 at the end of the chapter for details) simplifies the presentation of the information by presenting it as linear text that a user can 'cursor around'. Webbie also provides functionality especially aimed at bind users. For example, to be able to jump over the large number of links that appear at the beginning of many webpages. The second application adaptor sits between Microsoft Outlook and a screen reader. It provides access to information that is very difficult for a screen reader user to access in a straightforward way. It also provides additional functionality such as being able to search for free diary slots of specific durations subject to specific constraints (for example, 'find me a 2-h slot on a Friday between 12:30 and 17:30'). This type of functionality is not present in Microsoft Outlook because a sighted user can perform a visual search. Performing the same sort of search on the calendar using a screen reader is much more difficult. Further details have been given by Blenkhorn and Evans (2002).

13.10 Chapter Summary

The chapter introduced the basic definitions of screen magnifiers and screen readers, for these were the main two technologies described.

After a brief overview of the use of screen magnifiers, a detailed presentation of the different magnification modes was given. These included the mechanisms of Full screen magnification, Lens mode, fixed area magnification and some other magnification modes including strip magnification and multiple windows.

Screen magnifiers interact with other features of the visual interface and some aspects of mouse pointer behaviour, the smoothing of magnified text symbols to avoid aliasing, the problem of colour inversion and controlling the magnifier by means other than the mouse and keyboard were investigated. To approach the study of the architecture and implementation of a screen magnifier, the chapter presented a brief historical survey of screen magnifier technology before giving the description of a simplified architecture of a modern screen magnifier.

The screen reader approach to accessing the content of the visual interface involves audio description, explanation and control of the visual information displayed. Screen readers are also finding application with a number of telecommunications products and devices that use a screen interface, for example as when creating and managing text messages on mobile telephones.

The architecture and implementation of a screen reader is approached through a historical survey of early methods and this is followed by the presentation of the structure of the modern screen reader. The modern technology uses specialised techniques like hooking and the off-screen model (OSM) and these were described. The use of refreshable Braille displays was also described briefly at this point.

The chapter concluded with some material on hybrid screen reader-magnifiers, self-magnifying applications, self-voicing applications and application adaptors.

Questions

Q.1 Briefly, describe the purpose of a modern screen magnifier and its simplified architecture.

Q.2 What are the magnification modes: full screen mode, lens mode, fixed area mode and strip magnification?

Q.3 Briefly, describe the purpose of a modern screen reader and its simplified architecture.

Q.4 How does a typical screen reader track the focus? What is hooking?

Q.5 What is a refreshable Braille display? How is a Braille display used to achieve accessibility on a personal computer?

Q.6 Briefly describe the essential features of a hybrid screen magnifier, a self-magnifying application and a self-voicing application.

Projects

P.1 If you have access to a PC with Windows XP or Windows 2000, you should investigate the use of a screen magnifier. The Microsoft Magnifier can be run from Programs, Accessories, Accessibility, Magnifier or by pressing the Windows key and the ‘U’ key and starting the Magnifier.

As noted in the text, the Microsoft Magnifier is a special type of fixed area magnifier, which means that part of your computer screen will display the enlarged image and part the original image. The position of the enlarged image can be controlled by clicking and dragging its borders.

You should note that the Microsoft Magnifier does not track the text entry cursor in Word. In order to experiment with focus tracking in a word processor you are advised to use the Notepad application.

In order to get a feel for the use of a magnifier, it is suggested that you use Microsoft Magnifier to access the Web. We suggest that you work with the magnifier in the top part of the screen (as large as you can make it) and cover the bottom part of the screen so you cannot see the enlarged text. Try to access a number of websites. You should experience one of the chief difficulties faced by the user of a magnifier, *i.e.* you cannot see the full screen and have to spend a lot of time scrolling horizontally and vertically and have a good memory for information that has disappeared as you move around the screen. Note what happens whenever you move the mouse pointer. You should see that the enlarged image scrolls and that the mouse pointer is kept central. This issue is discussed in the text.

You should also try to use Microsoft Magnifier with Notepad. If the magnifier's options are set to track the keyboard focus and text editing (checkboxes on the Magnifier Settings form), you can get experience of what it is like to type a document using a magnifier. Because the magnification follows the cursor, it is relatively easy not to get lost whilst text is being typed. However, considerable amounts of scrolling will be required when text is to be read and edited.

Finally, you may like to check out the colour inversion options provided by Magnifier.

P.2 Get access to a Windows PC and run the Narrator screen reader (this can be found in Programs, Accessories, Accessibility, Narrator or by pressing the Windows key and the 'U' key and starting Narrator)—this should be present on a PC running Windows XP or Windows 2000.

First, with Narrator running, carry out some tasks interacting with the computer in your normal way (*i.e.* use the mouse), listening to what is spoken by the screen reader. You should find that in addition to echoing your typing and reading what is present on the screen, Narrator will also give information about the Windows element that has the focus. Try, for example, moving through a menu and listen to the description of elements as being 'menu item'. You should note that Narrator has limited functionality; it will not interact with the client area of Microsoft Word in the way that a full functionality screen reader would—for example, whilst characters are echoed, words are not spoken as you move the caret from line to line. If you wish to see how a screen reader works with a word processor, run Notepad. Narrator does not support the reading of the client area of a web browser so do not try surfing the Web—if you wish to do this see the project P.3.

Second, you should remember that a blind user does not have the option of using a mouse. To simulate this, try using the computer without using the mouse. To do this effectively you will need to use the keyboard operations of Windows (these can be found at http://support.microsoft.com/default.aspx?scid=kb;en-us;q126449; there are many other descriptions of Windows shortcuts and an Internet search on Windows Keyboard Shortcuts will yield many resources). This is a more accurate simulation of the experience of a blind user. However, you are still able to see the screen. Whilst you can still see the screen, run Notepad and type. Listen to what is spoken by the Narrator—you should hear characters being echoed. When you have typed a number of words on a number of lines, cursor around the client area (using the up, down, left and right arrow keys). You should find that as you move to a new line, the text on the line is spoken. As you move along a line from character to character, the character is spoken. As you move along a line from word to word (hold the Shift key at the same time as using the right and left arrow keys), you should find that the new word is spoken. This type of interaction is referred to as 'screen reading'.

Finally, turn the monitor off or shut your eyes. Use Narrator and the keyboard commands to interact with Windows. You may like to use Notepad to write up a short description of your experiences in using Narrator and send this to a printer.

In following this exercise, you should have noted the following.

When you 'cursored around' the Start Menu you may have noticed that the current speech was interrupted every time you pressed an arrow key, *i.e.* the screen reader does not finish what it is currently saying, but immediately gives an indication of the new focus. If this did not happen, the user could not be sure of the current focus and screen readers would be very frustrating to use.

Narrator gives more information than the text at the focus and this information is dependent on the type of element that has the focus. When interacting with menus, Narrator tells you the "Menu Item" details whichever cursor key you press (left, right, up or down).

In certain applications, the screen reader will change the speech dependent on the user's actions. For example, in Notepad a character is spoken if the user moves from character to character (using the right and left cursor keys), a word is spoken if the user moves from word to word (using Shift and the right and left cursor keys) and a whole line is spoken if the user moves from line-to-line (using the up and down cursor keys).

When new items appear on the screen, *e.g.* a new window, the user is informed of the title of the window and the element in that window that has the focus.

You should note that Narrator is a basic screen reader and not all applications interact with it in a way that is supported by other commercial screen readers. For example, it does not work when 'cursoring around' in Microsoft Word or Excel. However, Narrator provides a very convenient way of demonstrating the basic features of a screen reader.

P.3 You may like to see how the World Wide Web translates into text from a screen reader user. As noted in Project P.1, Narrator does not work very well with browsers. However, you can use a specialised web browser that is designed to make the Web simpler for screen reader users. To do this download and install Webbie from www.webbie.org.uk. This presents a website in text augmented by descriptions as to whether the text forms a link or an element of a form. Use Webbie and Narrator to access your favourite websites. This will give you an impression of not only how the Web can appear to a blind user, but also how a screen reader can be used to read text a line at a time—use the arrow keys to move the cursor through the text.

P.4 You may like to experience an application that provides self-voicing and self-magnifying. EdWord is a simple word processor that is designed to be used by people with sensory and cognitive impairments. EdWord can be downloaded from www.deafblind.org.uk/software#ee and may be freely used and distributed. After installing EdWord, you should compare its features to Microsoft Magnifier and Microsoft Narrator.

References

AiSqBi, undated, http://www.aisquared.com/Products/BigShot/BigShot.htm

Baude, A., Blenkhorn P., and Evans, G., 2001,The Architecture of a Windows 9x Full Screen Magnifier using DDI Hooking, In Assistive Technology – Added Value to the Quality of Life (AAATE 2001) Eds. C. Marinèek, C. Bühler, H. Knops, & R. Andrich, IOS Press, Amsterdam, pp 113–118

Blenkhorn, P. and Evans, G., 2001, The Architecture of a Windows Screen Reader, In Assitive Technology – Added Value to the Quality of Life (AAATE 2001) Eds. C. Marinèek, C. Bühler, H. Knops, and R. Andrich, ISBN 1383-813X, IOS Press, Amsterdam, pp 119–123

Blenkhorn, P., and Evans, G., 2002, Augmenting The User Interface Of Standard Applications For Improved Screen Reading For Blind People, Seventeenth Annual International Conference on Technology and Persons with Disabilities", Los Angeles, March

Blenkhorn, P., Evans, G., and Baude, A., 2002, Full Screen Magnification Using Direct X Overlays, IEEE Trans. on Neural Systems and Rehabilitation Engineering, Vol. 10, No 4, pp 225–232

Evans, G., and Blenkhorn, P., 2003, Architectures of Assistive Software Applications for Windows-based Computers, J. Computer and Network Applics, Vol. 26, pp 213–228

Fairweather, P.G., Hanson, V.L., Detweiler, S.R., and Schwerdtfeger, R.S., 2002, From Assistive Technology to a Web Accessibility Service, Proc. ASSETS'02, ACM Press, pp 4–8

FraAudio, undated, http://www.audiodata.de/e/about/index.php

FrSc, undated, www.freedomscientifc.com

Gibson M., and Reiser N., 1987, Lyon Large Print Program, Senus, pp 32–34, Summer

Griffith, J., 1987, DP-11, Sensus, pp 16–19, Summer

IBM, undated, http://www-306.ibm.com/able/solution_offerings/hpr.html

Kurniawan, S.H., King, A., Evans, D.G., and Blenkhorn, P., 2003 A User Evaluation of a Joystick- Operated Screen Magnifier, Proceedings CHI 03, ACM Press, pp. 25–32, April

MagM, undated, http://www.screenreader.co.uk/pages/magnice.htm

Marsh, K., 1993, Win 32 Hooks, Microsoft Developer Network, July 29

Microsoft, 1999, www.microsoft.com/Windows/platform/Innovate/Marketbulls/msaakybull.asp, current at December

Rubenstien, R., and Feldman, J., 1972, A controller for a Braille Terminal, Comm. ACM, Vol. 15, No. 9, pp 841–842, Sept.

SOS, undated, www.stonesoupsoftware.org

Spiry, B., 1987, Vista, Sensus, pp 35–37, Summer

Sterling, T.D., Lichstein, M., Scapino, F., and Stuebing, D., 1964, Professional Computer Work for the Blind, Comm ACM, Vol. 7, No. 4, April, pp 228–230, 251

WEB, undated, www.webbie.org.uk

14 Speech, Text and Braille Conversion Technology

Learning Objectives

Text in electronic form is a key and increasingly important intermediary in allowing access to information by visually impaired and blind people using assistive technology. Once text is in electronic form, it can be transmitted to distant recipients, read aloud using synthetic speech, converted to Braille media and displayed in large print for visually impaired readers. Text can be produced in electronic form using input from a keyboard, speech and/or Braille. Conversion technology is thus the enabler or intermediary which makes possible the various input and output combinations that allow blind and visually impaired, as well as other disabled people to access electronic text. The key conversion technologies of speech-to-text (STT), text-to-speech (TTS), Braille-to-text (BTT) and text-to-Braille (TTB) are the focus of this chapter, which has the following learning objectives:

- Gaining an in-depth understanding of the fundamental scientific principles that support spoken language technology.
- Learning some of the details of speech-to-text and text-to-speech conversion.
- Understanding the basic concepts of Braille conversion technologies.
- Gaining an appreciation of the application of these conversion technologies to assistive technology systems for visually impaired and blind people.
- Learning about the benefits and limitations of the current state-of-the art technologies for conversion applications.

14.1 Introduction

14.1.1 Introducing Mode Conversion

Human communication is multimodal. People use a number of different types of communication, including images, text, gestures, oral speech, sign languages, touch, mime, body language, facial expressions and music, to communicate with each other. Two or more means of communication are frequently combined to improve comprehension. For instance, in face-to-face communication, speech is often combined with gestures and mime. In addition, most people are able to switch between different types of communication signals, for instance, from saying "please could you give me that loaf of bread" to actually pointing at the item desired. Each communication signal is received and analysed by one of the senses. This fact can be used to categorise the various communication signals into modes (acoustical, visual, and tactile). Blind people have no or only very limited access to visual communication modes. Therefore technical support is required to present information which was originally intended to be received by the visual mode into other modes, including the following.

Speech

Linguistic information, which is very important for communication, can be presented in either written or spoken form. Only the spoken form is accessible to blind people and therefore the conversion of text to speech (and *vice versa*) is a key procedure for assistive technology. The fundamental differences between written and spoken language increase the complexity of the conversion technology. Therefore, spoken language technology (SLT) is a demanding and still evolving field of research and development. The state of the art in this field (which is still unsatisfactory, particularly when compared to human speech output), will form the main topic of this chapter.

Other audio information

The visual channel is continuously exposed to a wide range of non-linguistic information, selected domains of which, can be converted into audio information using assistive technology. A typical example is the transformation of document structures to audio information. In such cases, there are benefits in complementing speech information by additional audio information. The conversion technology involved is of increasing importance, as it is the key for blind people (but not deafblind people) to access the World Wide Web successfully.

Tactile information

Some information in the visual channel can be successfully converted into tactile information. Maps and diagrams are typical examples, as they can be equipped with contours which can be felt with the fingertips. However, there are issues of the appropriate amount and type of detail and tactile diagrams are discussed in more detail in Chapter 4. Text can be represented in tactile form using text-Braille conversion technology. Braille is well-established, though computer programmes for converting text to Braille are considerably more recent. Learning to read Braille requires training, time and effort (similarly to learning to read text) and is difficult for older people. Braille is increasingly being replaced by speech-based methods. However, these methods are not appropriate for deafblind people and there would be benefits to blind people in being able to choose whether to access text through Braille or speech.

Blind people require a number of different converters in order to be able to access visual information in their preferred format. This chapter will focus on the conversion tasks that are closely related to linguistic information, namely speech-to-text, text-to-speech, and text-to-Braille.

14.1.2 Outline of the Chapter

The main aim of this chapter is to provide an understanding of spoken language technology (SLT) and its applications to support visually impaired and blind people. Although SLT is still not very widespread, it has a very large potential for future development. However, this very practical goal is supported by underlying theory and therefore much of the contents of the chapter are necessarily theoretical in order to provide the background for understanding the practical applications. SLT is complicated and the currently available equipment has still to be perfected, resulting in problems and unsatisfactory performance at times. A good understanding of SLT fundamentals is required to recognize the reasons for these problems as well as the potential for future development of SLT.

The technical presentation of the chapter begins with Section 14.2 which provides the reader with the prerequisites necessary for understanding speech and language related technology. This includes aspects of signal processing and, in particular, spectral analysis (Section 14.2.1), as well as some aspects of linguistics (Section 14.2.2). This leads to a general scheme for a speech processing system, shown in Figure 14.4 which will serve as a didactic framework for the discussion presented in the rest of the chapter.

A detailed explanation of speech-to-text conversion is given next, in Section 14.3, as this is required to assess the capabilities and the problems of speech recognition equipment. A presentation of selected principles of pattern recognition in general is given in Section 14.3.1, and speech recognition, in particular is described in Section 14.3.2. Selected classes of speech recognizers are described in Section 14.3.3.

A key technology in assistive technology and rehabilitation engineering is text-to-speech (TTS) conversion and a detailed explanation of this technology is presented in Section 14.4. The principles of speech production in general are given in Section 14.4.1, with particular attention to the audio output (Section 14.4.2). Finally, an overview of the existing classes of TTS systems is presented (Section 14.4.3).

The basic principles of Braille conversion are introduced in Section 14.5 and, since Braille technology only appears at the level of written language (text level), this is a short section. It is followed by Section 14.6 that looks at the application of different conversion technologies to commercial equipment for blind and visually impaired people. This is a very large field, which it is partially covered by other chapters in this book. Therefore, this section is restricted to summarising the technology that is considered to be of most practical use.

Finally, in Section 14.7, some open problems and their potential solutions are discussed. This section commences with a few remarks on the current state of the art (Section 14.7.1). A number of problems are identified (Section 14.7.2) and suggestions for future directions for research and development with a view to resolving these problems are given (Section 14.7.3).

The following two remarks refer to topics which will not be covered in this chapter:

- Readers require some understanding of the fundamentals of sound and hearing, in order to understand speech technology. These topics are not presented here, as they are covered in the first chapter of the previous volume of this AT book series (Hersh and Johnson 2003), to which readers are referred.
- The algorithms used in speech signal processing will not be discussed. The interested reader is referred to standard textbooks, such as those by Deller *et al.* (1993) and Rabiner and Juang (1993).

14.2 Prerequisites for Speech and Text Conversion Technology

14.2.1 The Spectral Structure of Speech

From the speech signal to the spectrogram

From the physical point of view, the speech waveform is a function of sound pressure against time. The speech signal can be recorded and processed using a microphone, lowpass filter, analog to digital converter, and a sample and hold circuit to give a digital representation of the signal in terms of a sequence of discrete measured values called samples. Good speech quality can be obtained with a bandwidth of 8 kHz. (This is greater than the bandwidth of telephone speech which is between 300 Hz and 3400 Hz.) From the sampling theorem of Kotelnikov and Shannon (Shannon 1949), a sampling frequency of at least 16 kHz is required, giving a time interval of 62.5 μs between two neighbouring samples.

The following steps will be illustrated by means of an example. Figure 14.1a shows the waveform of the word *Amplitude*. (It was pronounced by a male speaker in German, but the example is language independent.) To give a feel for the quantity of data involved in speech processing, it should be noted that this relatively short word which represents 1.3 s of speech requires $(1.3\,\mathrm{s})/(62.5\,\mu\mathrm{s}) = 20,800$ samples for a sampling frequency of 16 kHz, which only just satisfies the sampling theorem. Each sample requires 2 bytes of storage to ensure sufficient accuracy.

The human inner ear acts as a spectrum analyser. Consequently, it is useful for technical systems to produce a spectral representation of the speech signal. A spectrum describes the composition of a signal from simple (harmonic) signals at particular frequencies, that is, it is a representation of amplitude against frequency. The required transformations are well known in signal processing and are summarized in Table 14.1. The relevant formulae can be found in most textbooks on signal processing.

However, speech is not a stationary signal. Therefore analysis should be based on segments of the signal (called windows) which can be considered to be "quasi stationary". A window can be considered to be an "analysis period" and therefore

Table 14.1. Overview of the different spectral transforms and the properties of their spectra

	Time-continuous signals	Time-discrete signals
Periodic signals	**Fourier series**	**Discrete Fourier transform (DFT); special version: Fast Fourier transform (FFT)**
	Non-periodic line spectrum	Periodic line spectrum
Non-periodic signals	**Fourier transform (Fourier integral)**	**Discrete time Fourier Transform (DTFT)**
	Non-periodic continuous spectrum	Periodic continuous spectrum

the upper row and the rightmost column of Table 14.1 gives the discrete Fourier Transform (DFT) as the appropriate transformation to be applied.

The length of the window plays an important role, so that the longer the window, the more details that can be identified in the spectrum. In the example, this can be observed by comparing the pictures in Figure 14.1b,c. Figure 14.1b has been calculated for a longer window and therefore provides more spectral information. On the other hand, a shorter window allows a better localisation of the spectrum on the time axis. Therefore, choice of an appropriate window length requires trade-offs between detailed spectral information and precise localisation, as well as the further practical constraint of the number of samples per window to be a power of two. This condition is required by the fast Fourier transform (FFT), which is an efficient algorithm for calculating the DFT. A choice of 256 samples, which corresponds to a window of 16 ms, is a good compromise.

This process results in the analysis of a single window. The complete characterisation of a speech signal requires the window to be shifted in short time steps. This results in a sequence of separate short time spectra which is represented graphically in Figure 14.1d. The amplitude of the spectrum is plotted in the time-frequency plane. This results in essentially the same input information as is available to a speech recognizer. The resulting graph, such as the "waterfall" shown in Figure 14.1d is not easy to interpret for a human observer. An easier to understand visual representation can be obtained in the form of a quasi-geographical map with the amplitude of the spectrum either coded in grey scale or represented by different colours. This is illustrated in Figure 14.2.

The maps which are produced in this way are called spectrograms. What the resulting spectrogram looks like will generally be influenced by the window length. In particular a longer window, resulting in greater spectral detail, as shown in Figure 14.1b, will give a *narrowband spectrogram* of the type shown in Figure 14.2a, whereas a shorter window, resulting in better time resolution, but less spectral detail as shown in Figure 14.1c will give a *broadband spectrogram* of the type shown in Figure 14.2b.

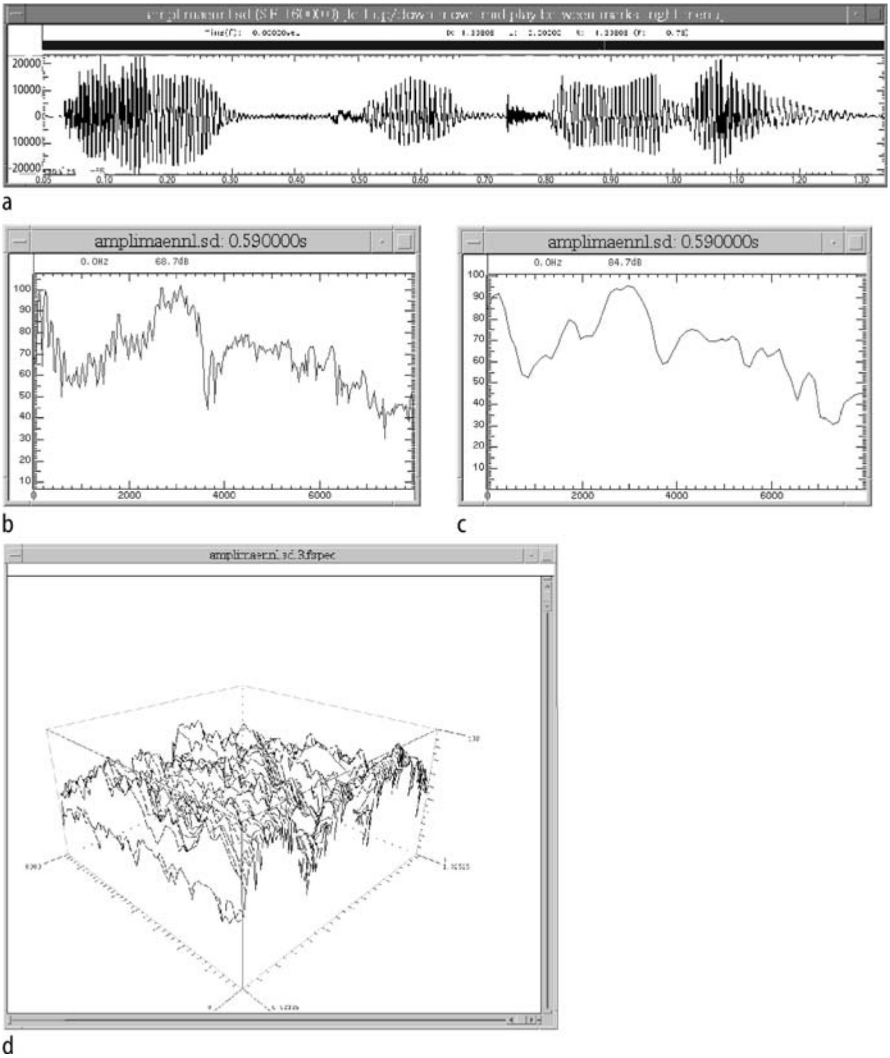
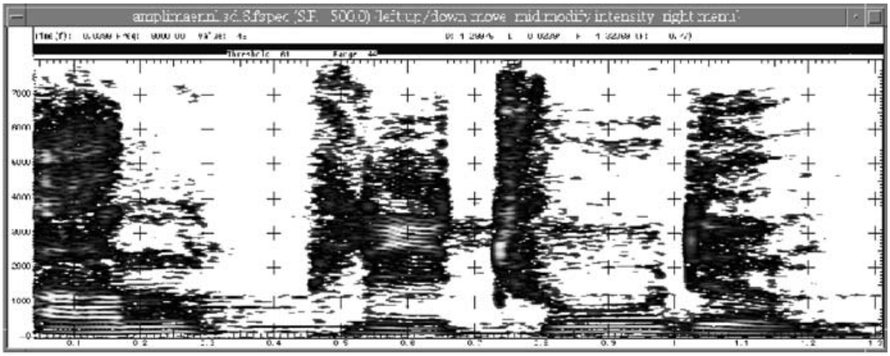
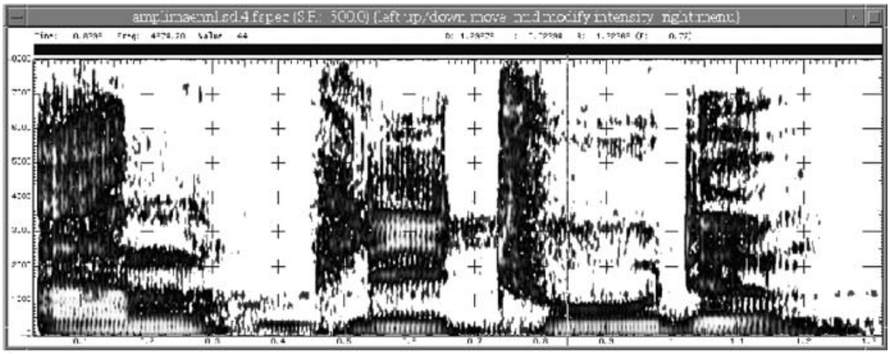


Figure 14.1a–d. Example, showing the way from the speech signal to the spectrogram: **a** acoustical waveform (sound pressure *vs* time) of the word "Amplitude" pronounced in German by a male speaker; **b** spectrum of the sound [i] of the given word, calculated by fast Fourier transform (FFT) of a speech segment (window) of 32 ms. For such a "long" window, the spectral details can be observed very well (narrowband spectrum); **c** spectrum of the same sound [i], calculated from a window of merely 8 ms. In this case, a better presentation of the spectral envelope is obtained (broadband spectrum); **d** if the complete word is analysed window by window, we obtain a sequence of spectra according to **b** or **c**, respectively, which form a relief from mountains and valleys over the time-frequency plane



a



b

Figure 14.2. Visualization of the sequence of spectra by means of a spectrogram (continued from the example presented in Figure 14.1). Because the three-dimensional presentation from Figure 14.1d is hard to interpret, a map-like presentation as a top view of the spectral "landscape" is preferred, called spectrogram. In a spectrogram, the abscissa acts as the time axis, the ordinate as the frequency axis, and the spectral magnitude is coded in colours or in a grey scale: **a** narrowband spectrogram of our example word "Amplitude", composed from spectra like Figure 14.1b; **b** broadband spectrogram of the same word, composed from spectra like Figure 14.1c

Excitation source and articulation tract

There are a number of different types of speech sounds, which are produced in slightly different ways. One of the main distinctions is between voiced and unvoiced sounds. *Voiced* sounds are produced by a process called *phonation* in which an air stream is conducted from the lungs through the larynx and leads to a quasi-periodic opening and closing of the vocal cords. The resulting speech signal is quasi-periodic. *Unvoiced sounds* are produced without phonation.

In the spectrograms of Figure 14.2 voiced sounds are clearly apparent, whereas unvoiced sounds are not very distinctive. In particular, voiced sounds show clear regularities or periodicities either in the frequency direction of the narrowband spectrogram, or in the time direction of the broadband spectrogram. These pe-

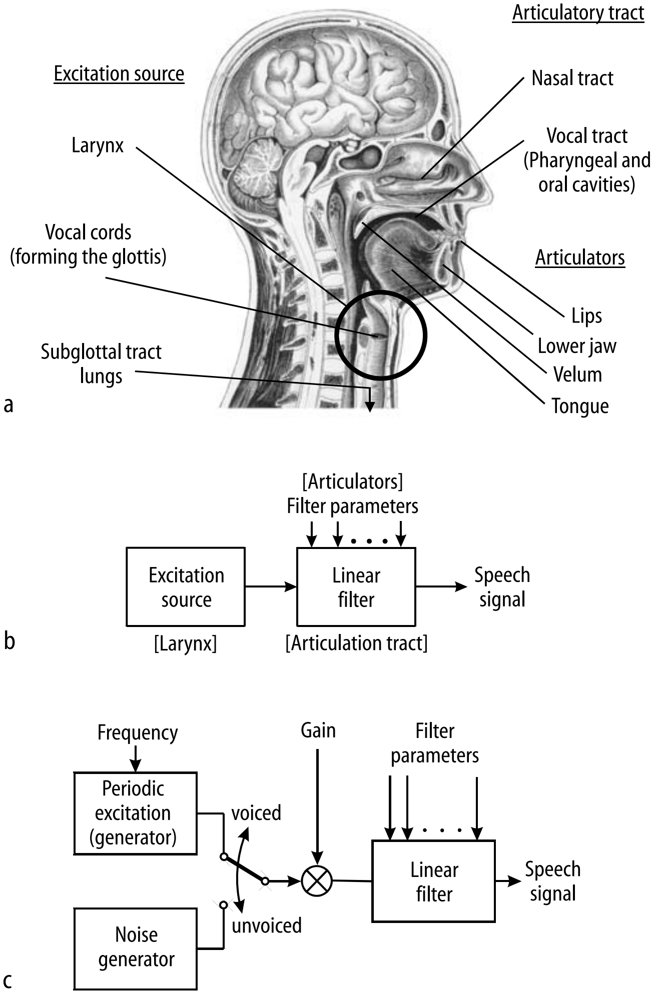


Figure 14.3a–c. The articulation of sounds and the linear model of it: **a** articulatory organs of the human; **b** linear model of the production of voiced sounds; **c** block diagram of a parametric speech synthesis system basing on the linear approach of **b**

riodicities reflect the periodic excitation of the larynx which produces voiced sounds.

Figure 14.3a illustrates the fact that the vibrations of the vocal cords of the larynx are influenced by the shape of the different cavities of the *articulation tract* that they pass through on their way from the larynx to mouth. The resulting modulation of speech allows the speaker to position the lips, tongue and other articulatory organs appropriately to produce the desired sound. Fortunately the process is automatic for people speaking their native language. However, it is generally less easy for people trying to produce the correct sounds in a foreign language.

A linear model can be used to describe a static "snapshot" of the complicated movements of the articulatory organs (Fant 1960). This model requires the assumption that the excitation source and the articulation tract can be modelled independently of each other. Figure 14.3b shows that the larynx can be modelled by a source of a periodic signal, the frequency of which can be controlled. Let $X(\omega)$ be its spectrum. This signal acts as the input to a linear system (or filter) which models the articulation tract. Let $H(\omega)$ be its transfer function. Then the spectrum $Y(\omega)$ of the signal produced at the filter output is given by

$$Y(\omega) = H(\omega) \cdot X(\omega) . \qquad (14.1)$$

The positions of the articulatory organs are coded in the filter parameters. The system can be used as a "parametric" speech synthesis system in combination with known values of the sequence of filter parameters. The system should include control of the excitation frequency to give the "speech melody" or *intonation*. It should be noticed that this system can only produce voiced, but not unvoiced sounds, as it is based on phonation. A noise generator is required to produce unvoiced sounds, as they do not have a periodic structure. This results in the configuration shown in Figure 14.3c which can be really used to synthesize speech.

It should further be noted that the linear model is also important in speech recognition. As described above, speech sounds are largely characterized by their spectral shape. Therefore, the spectrum of the speech signal provides an appropriate input for a speech recognizer. However, if possible, the influence of the excitation source should be removed from this spectrum, as the source signal contains additional information, such as the speaker's gender, which is not required by the recognition process. In terms of Equation 14.1, this means that the desired spectrum $H(\omega)$ needs to be extracted from the measured spectrum $Y(\omega)$. Since multiplication in the frequency domain corresponds to convolution in the time domain, the required operation is called *deconvolution*. In practice, this can be performed using the so-called *cepstrum*. The cepstrum is calculated by taking the logarithm of the spectrum and then the inverse Fourier transform. Examination of the resulting cepstral coefficients shows that the coefficients describing the excitation can be separated from those describing the influence of the vocal tract. Alternatively, $H(\omega)$ can be estimated by means of the method of *linear prediction coding* (LPC).

14.2.2 The Hierarchical Structure of Spoken Language

Spectral properties of the speech signal were discussed in the previous section. However, the spectral description only refers to the physical or acoustical properties of speech, whereas one of the most important properties of speech is its use to intentionally transmit messages, coded by following the rules of the language being spoken, that is, it is *spoken language*. This means that units of speech, such as syllables, words, phrases and sentences can be identified in addition to the basic speech sounds. These units are combined according to rules rather than in an arbitrary manner. For example, words are combined to make phrases and/or sentences according to grammatical rules. Consequently there is a *hierarchy* of levels. At each level, a system of rules or constraints determines how the units at this level are composed from units at the level below. The importance of this hierarchical structure is emphasised in Levinson's classical paper (Levinson 1985): "The constraints build redundancy into the code, thereby making it robust to errors caused by ambient noise or made by the speaker. Thus a relatively few primitive patterns can be combined in a multilevel hierarchy according to a complex code to form a rich, robust information-bearing structure." This has the following consequences for engineering speech processing solutions:

- Speech processing systems should also have a hierarchical organisation, with the number of levels determined by the aims of the system. A recognizer for single commands only requires the hierarchy from the signal to word level, whereas a system for translating spoken language has to model the complete hierarchy up to the semantic level.
- A set of rules, constraints or decision rules is required at each level. Therefore, the system should include a set of knowledge bases which can be implemented in several different ways. A number of examples of knowledge bases which are used in real systems are summarized in Table 14.2.

Table 14.2 has separate columns for speech recognition and speech synthesis systems, due to their different requirements and different history. However, there has been a convergence of the knowledge bases for speech recognizers and speech synthesizers over the last decade. It is this convergence in particular which allows a unified approach to the representation of the components of the two types of systems.

Table 14.2. Databases for speech recognizers and synthesizers

	Recognizer	Synthesizer
Sentence level	Language model Stochastic n-gram grammar Regular (automata) grammar	Rules for detecting phrase boundaries and accents In special cases: language model
Word level	Pronunciation dictionary (maybe with pronunciation variants)	Rules for grapheme-to-phoneme conversion with exception dictionary (no sharp boundaries to a complete pronunciation dictionary)
Sound level	Phoneme models (monophones, triphones, *etc.*) mostly as stochastic models (hidden Markov models = HMM, stochastic Markov graphs = SMG)	Sound or allophone models or combinations of them (monophones, diphones, syllables, *etc.*) as waveforms (predominant), parametric representations (*e.g.* formant models, LPC models), stochastic models (*e.g.* HMM)
Acoustical or physical level	Feature vectors basing on spectral or related representations	Feature vector for controlling a synthesis filter, if present, otherwise simply samples

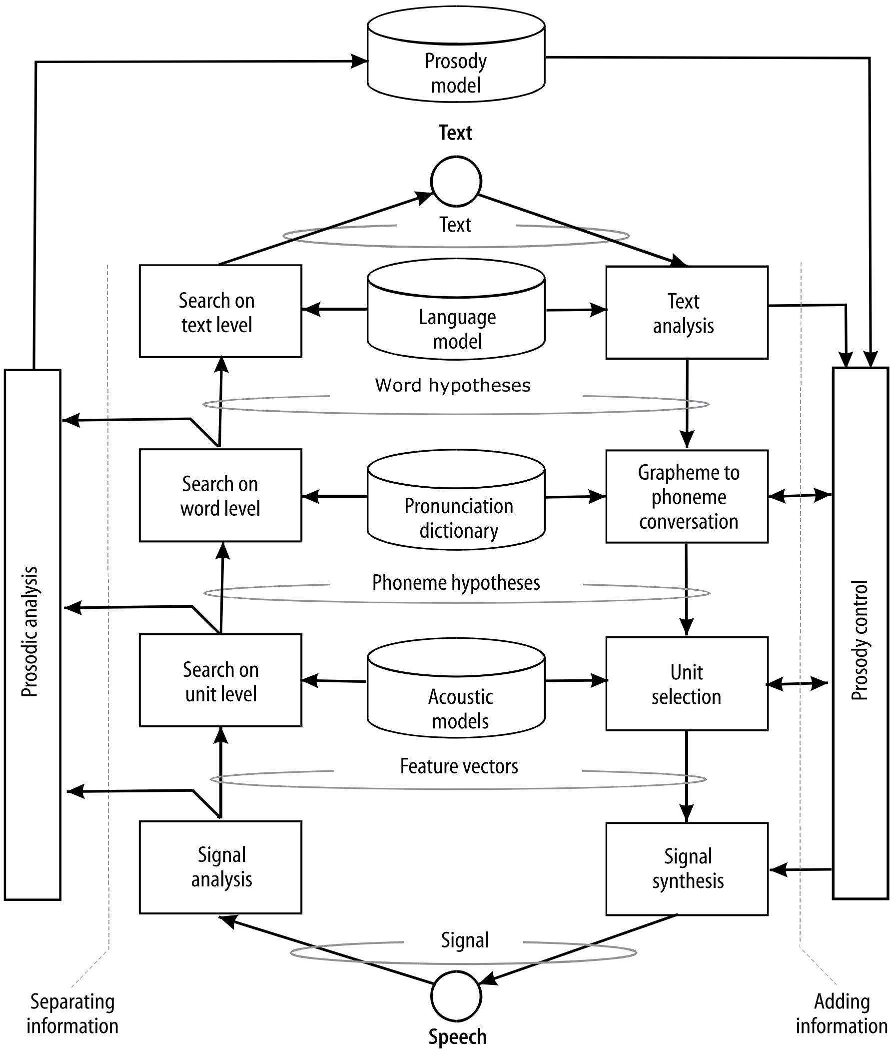


Figure 14.4. Unified approach for speech synthesis and recognition (UASR) as a prototype of a speech processing system (Eichner *et al.* 2000)

Figure 14.4 illustrates the use of a set of common databases to support both an analysis or recognition branch (left) and a synthesis branch (right). This general scheme for a speech processing system will be used to develop the block diagrams of more specialized systems in the subsequent sections. The figure presents the levels up to the text or sentence level, as this corresponds to the requirements of most of the current applications.

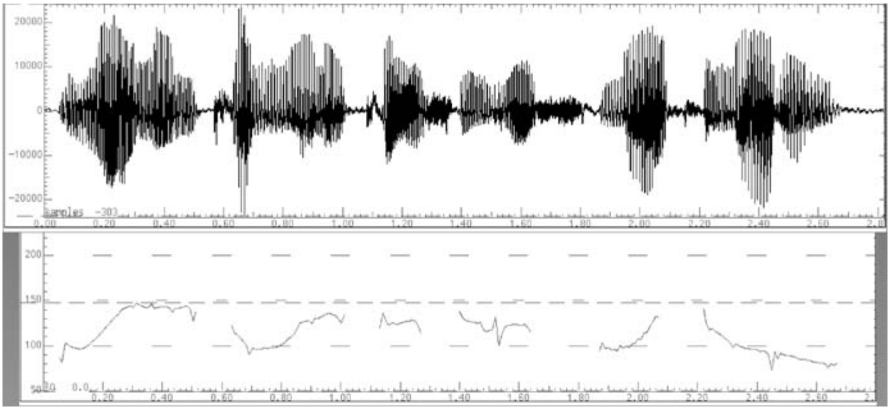


Figure 14.5. Example of an intonation contour produced by a male speaker. The *upper part* shows the waveform of the sentence *"Human communication is multimodal." Below*, the pitch of the voiced parts of this utterance is plotted *vs* time

14.2.3 Prosody

Phonetics defines a set of distinctive features which are used to characterize the different speech sounds or phones. For instance, the distinction between voiced and unvoiced sounds has already been mentioned. However, there are also a number of features or physical parameters of speech which are not used to distinguish between different phonemes (at least in European languages). For example, the sound "a" remains "a" whether spoken with a low fundamental frequency (male voice) or a high fundamental frequency (female voice). This group of parameters is important, although these parameters do not contribute to the description of speech sounds. The group includes the following parameters:

- The pitch or fundamental frequency (which gives the intonation, as illustrated in Figure 14.5).
- Sound duration.
- Sound energy.

The contours of the variation of these three parameters with time together form the phenomenon called *prosody*. Investigation of these contours in spoken language, for instance the intonation contours of a statement and a question, shows that the significant time periods are larger than the elementary segments or phones. Therefore, prosody affects the *suprasegmental* as well as the segmental features of speech. Table 14.3 summarizes the different functions of prosody. Since prosody is very important in speech communication, it needs to be considered in electronic speech communication systems. From the suprasegmental characteristics of prosody it follows that the related components "embrace" the segmental oriented branches of analysis and synthesis as shown in Figure 14.4. In the analysis branch, the prosodic parameters are extracted from the speech signal and interpreted to

Table 14.3. Functions of prosody based on a classification by Fujisaki (1997)

Linguistic information	Paralinguistic information	Nonlinguistic information
Transfers the linguistic structure of an utterance	Transfers additional information	Transfers information about the speaker
• Groups the information by phrasing. • Emphasizes the information by accentuation. • Determines the type of a sentence (*e.g.* question).	• Modifies or supplements the linguistic information. • Clarifies the contents in case of ambiguities.	• Age. • Gender. • Physical conditions. • Emotional state.

improve the recognition results. In the synthesis branch, prosody is added to the segments to make the synthesized speech more natural and acceptable.

However, in practice prosodic information has only been applied to speech recognition systems in a few cases, such as the translation system Verbmobil (Wahlster 2000). In contrast, speech synthesis systems generally have a prosodic component, as the acceptance of synthetic speech strongly depends on its naturalness, which is significantly improved by good prosody. Therefore, modelling prosody has been an important research topic for a number of years. However, the results of this research have not been completely satisfactory due to the complex and important role of prosody, as illustrated in Table 14.3.

14.3 Speech-to-text Conversion

14.3.1 Principles of Pattern Recognition

Speech recognition includes interactions between a number of processing steps at different levels of the speech and language hierarchy. The lowest (acoustic) level is particularly important, as it is required by both simple and sophisticated speech recognition systems. To support understanding of the fundamentals of speech recognition, a brief overview of the theory of pattern recognition will be presented.

A pattern recognition system consists of two subsystems, an analyser and a classifier, as illustrated in Figure 14.6. As discussed in Section 14.2.1, the analyser produces a spectral representation of the speech signal. This leads to a multidimensional description of the input by means of so-called *feature vectors*. It should be noted that feature vectors are generally used in pattern recognition systems and not only in pattern recognition systems for speech recognition.

The feature vectors form the input to the classifier, which has to decide which class k the particular input belongs to. Two different types of classification are carried out using different types of input. Numerical classification is based on a single feature vector $\boldsymbol{x}$. In this case each feature vector is assigned to a class.

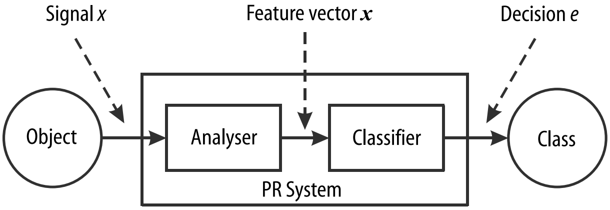


Figure 14.6. Structure of a pattern recognition system

Structural classification is based on a sequence $\underline{\boldsymbol{x}}$ of feature vectors. Numerical classification will be considered first and used to introduce some basic concepts in pattern recognition.

Numerical classification

In numerical classification the set of feature vectors $\boldsymbol{x}$ produced by the analyser is mapped to the set of available classes k. The set of all feature vectors (together with the required operations) can be considered to be a vector space (the so-called feature space) and therefore classification simply corresponds to a partitioning of the feature space. This will be illustrated by means of an example of the representation of a typical vowel spectrum in Figure 14.7.

Feature vectors are generally of high dimension. However, to simplify the example, a low dimension feature vector from speech processing will be used. As shown in Figure 14.1c, the spectral envelope of a vowel is characterized by a certain combination of maxima and minima. It can be shown that long (static) vowels can be satisfactorily described by the frequencies of the first two maxima. These frequencies are called the first and second formant frequencies, F_1 and F_2, and are shown in Figure 14.7. Therefore, the example is based on a two-dimensional feature space with feature vectors:

$$\boldsymbol{x} = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} . \tag{14.2}$$

A set of classified vectors (training data set) is required during the teaching phase to calculate the correct partitioning of the feature space. In the example, the

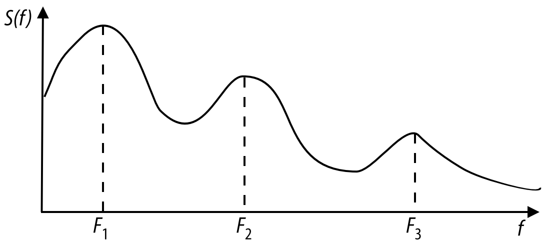


Figure 14.7. Typical formant structure in the envelope of the power density spectrum of a long vowel (generalised from real spectra like Figure 14.1c)

formant frequencies of five German vowels pronounced by 31 speakers were measured, giving five classes, one corresponding to each vowel. The training set has $5 \times 31 = 155$ elements and their positions are shown in Figure 14.8a. It can be easily generalised from the example that the training elements of a certain class form a cluster in the feature space. The clusters of different classes are more or less separable. Training of the classifier requires partitioning the feature space by calculating appropriate boundaries between the clusters. Thus the feature space is partitioned in the training phase and subsequent recognition of a feature vector representing an unknown object only requires the classifier to determine which partition of the feature space the feature vector belongs to.

There are a number of different methods for partitioning the feature space, of which three will now be discussed briefly.

- *Linear classifier*. A linear classifier separates the classes by planes (or lines in the two-dimensional example). It can be implemented easily, but its performance is limited. Figure 14.8b illustrates the simplest approach, with the boundaries shown as the perpendicular bisectors between the centre vectors of pairs of clusters from the training set (*distance classifier*).
- *Bayesian classifier*. The performance of the classifier can be essentially improved by using statistical information estimated from the training database. This includes the likelihood $P(\boldsymbol{x}|k)$ or the corresponding probability density function (pdf) $p(\boldsymbol{x}|k)$ and the *a priori* probabilities $P(k)$ for all classes k. Bayes' rule can then be used to give the *a posteriori* probabilities which are required for the classification, namely

$$P(k|\boldsymbol{x}) = \frac{P(\boldsymbol{x}|k) \cdot P(k)}{P(\boldsymbol{x})}. \tag{14.3}$$

  If the pdf $p(\boldsymbol{x}|k)$ is estimated as a Gaussian distribution, the boundaries between the classes in the feature space are second order, as can be seen in the vowel example in Figure 14.8c.
- *Neural classifier*. Sets of elementary subsystems called neurons can be used in the classification. The combination of the non-linear transfer functions of the neurons gives a non-linear input-output mapping of the whole network which can be learned during a training phase by tuning the weights of the neurons. As shown in Figure 14.8d, a neural classifier uses non-linear boundaries to separate the classes and can be very successful.

The different types of classifiers learn different sets of model information $\mathcal{M}_{\mathrm{k}}$ for the classes $k = 1, \ldots, K$ during the training phase. However, the different classifiers use the same classification principle after the training phase. For a given feature vector $\boldsymbol{x}$, the values of a discrimination function $d_{\mathrm{k}}(\boldsymbol{x})$ are calculated for all the classes k. The function $d_{\mathrm{k}}(\boldsymbol{x})$ gives a measure of the similarity between the feature vector $\boldsymbol{x}$ and the model information $\mathcal{M}_{\mathrm{k}}$. Therefore, choice of the best value of $d_{\mathrm{k}}(\boldsymbol{x})$ will give the class e which the classifier should select as

$$e = \arg \operatorname*{ext}_{k=1,\ldots,K} d_{\mathrm{k}}(\boldsymbol{x}). \tag{14.4}$$

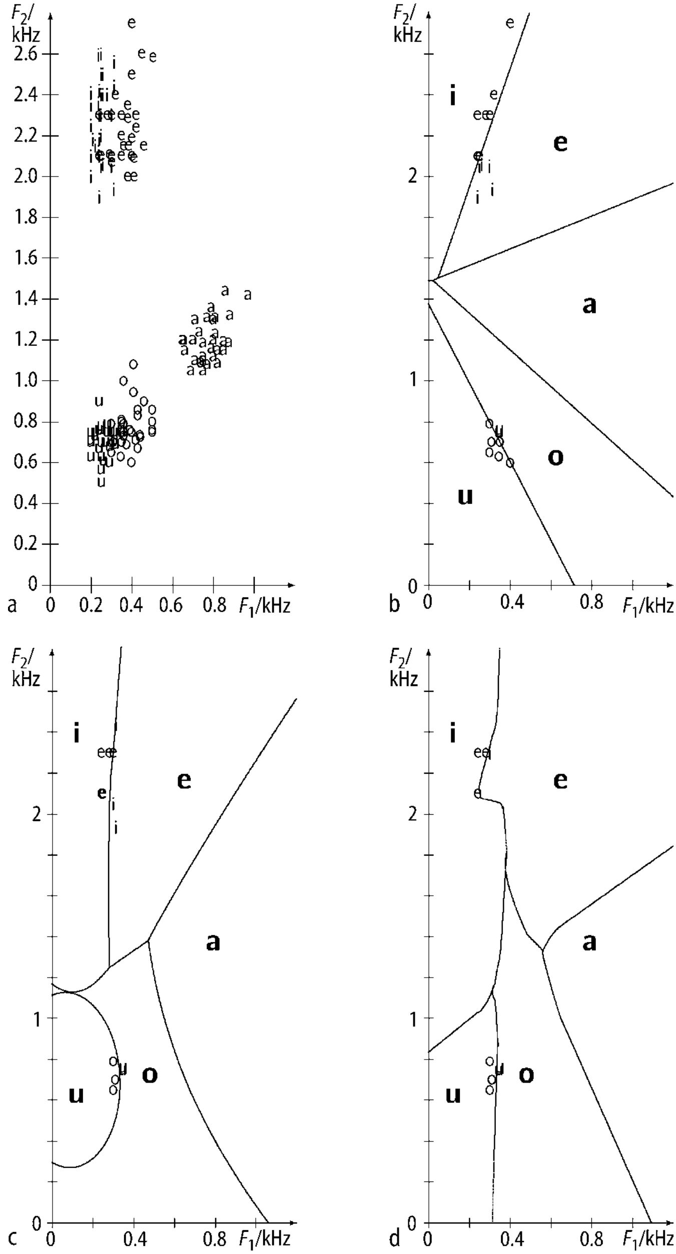


Figure 14.8a–d. Example for classification (Reproduced from *Signalanalyse und -erkennung*, 1998, page 290 (14.8a), 303 (14.8b), 313 (14.8c) and 367 (14.8d), Hoffman, R., © Springer-Verlag. Used with permission.): **a** the feature space for vowels described by their first formants F_1 and F_2 as explained in the text. The elements of the training set are indicated by small letters. In the other figures, the partitioning of the feature space is shown for the distance classifier (**b**), the Bayesian classifier with Gaussian pdfs (**c**), and a neural classifier (**d**). In **b–d**, *large letters* describe the areas of the five classes, while *small letters* indicate those elements from the training set in **a** which are misclassified after determining the class boundaries

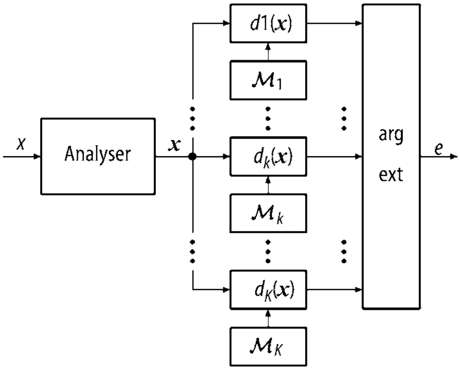


Figure 14.9. Structure of a recognition system. This structure is developed from the PR system Figure 14.6 considering the details resulting from the concept of discrimination functions (Reproduced from *Signalanalyse und -erkennung*, 1998, page 294, Hoffman, R., © Springer-Verlag. Used with permission.)

Equation 14.4 is called a decision rule and this general scheme is shown in Figure 14.9.

Different decision rules can be specified for different types of classifiers. For the distance classifier, $d_{\mathrm{k}}(\boldsymbol{x})$ is simply the distance $d(\boldsymbol{r}_{\mathrm{k}}, \boldsymbol{x})$ between a typical representative $\boldsymbol{r}_{\mathrm{k}}$ of class k and the feature vector, $\boldsymbol{x}$. The shortest distance corresponds to the best value of the discrimination function, and the decision rule is

$$e = \arg \min_{k=1,\ldots,K} d(\boldsymbol{r}_{\mathrm{k}}, \boldsymbol{x}) . \tag{14.5}$$

For statistical classifiers, $d_{\mathrm{k}}(\boldsymbol{x})$ is proportional to the *a posteriori* probability $P(k|\boldsymbol{x})$. The highest probability is the best value for the discrimination function, and the decision rule is

$$e = \arg \max_{k=1,\ldots,K} P(k|\boldsymbol{x}) . \tag{14.6}$$

The discussion so far has concerned the classification of patterns described by single feature vectors (numerical classification). However, speech signals are described by sequences of (spectral) feature vectors and therefore the calculation of distances or probabilities needs to be generalized to sequences of feature vectors. The same approach can be used as for a single feature vector with the discrimination functions $d_{\mathrm{k}}(\boldsymbol{x})$of the single feature vector $\boldsymbol{x}$ in Figure 14.9 and Equation 14.4 replaced by the functions $d_{\mathrm{k}}(\underline{\boldsymbol{x}})$of the sequence of feature vectors $\underline{\boldsymbol{x}}$. The classification of sequences of feature vectors was first used in speech technology during the 1970s (Levinson 1985).

Structural classification I: Dynamic time warping

The simplest approach is based on the adaptation of the concept of the distance classifier. This involves comparison of the sequence $\underline{\boldsymbol{x}}$ of feature vectors describing an unknown utterance with sequences $\underline{\boldsymbol{r}}_{\mathrm{k}}$ of feature vectors describing typical

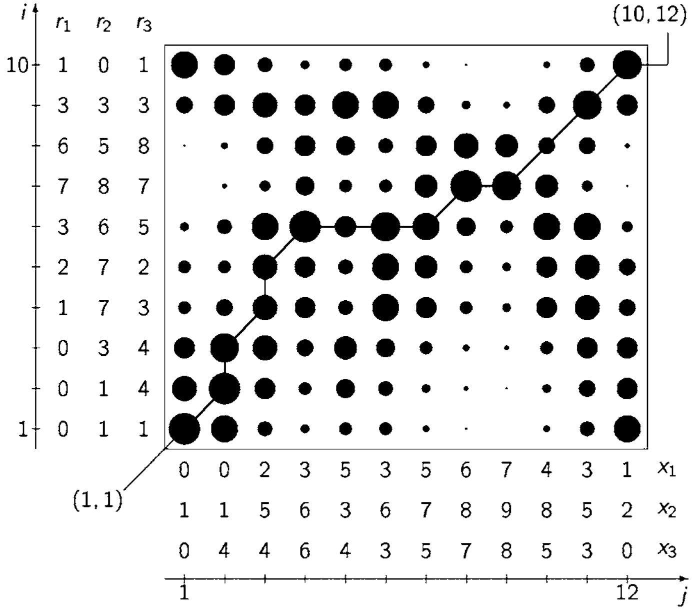


Figure 14.10. Example for the DTW algorithm (Reproduced from *Signalanalyse und -erkennung*, 1998, page 391, Hoffman, R., © Springer-Verlag. Used with permission.)

(reference) instances which represent the different classes, k. It is then necessary to solve the problem of how a "distance" between the two sequences of feature vectors can be computed. The approach will be explained by means of the example in Figure 14.10. For simplicity, the example only considers two sequences of feature vectors having three dimensions, whereas real systems involve vectors with 20 or more dimensions.

The method first requires all the distances between pairs of feature vectors to be calculated. This gives a matrix of local distances, where $d(\underline{r}_k(i), \underline{x}(j))$ represents the distance between the ith feature vector in the first sequence and the jth feature vector in the second sequence. In Figure 14.10, the local distances in the matrix are represented by black circles, with large circles indicating a low local distance or high similarity.

The next step involves the calculation of the global distance $D(\underline{r}_k, \underline{x})$ as the sum of all local distances on an optimal path through the matrix. This requires choice of the optimal path using a standard dynamic programming algorithm. This results in a modification of the decision rule (Equation 14.5) to replace the local distance by the global distance in the discrimination function:

$$e = \arg \min_{k=1,\ldots,K} D(\underline{r}_k, \underline{x}) . \tag{14.7}$$

This method is known as *dynamic time warping* (DTW), as dynamic programming is applied to sequences ordered in time.

Structural classification II: Statistical methods

Other and more powerful classification approaches, which will now be considered, are based on statistics. The decision rule (Equation 14.6) is now modified to

$$e = \arg \max_{k=1,\ldots,K} P(k|\underline{\boldsymbol{x}}) \tag{14.8}$$

where the sequences $\underline{\boldsymbol{x}}$ represent the speech segments. The *a posteriori* probability $P(k|\underline{\boldsymbol{x}})$ is calculated from the Bayesian rule according to Equation 14.3 as

$$P(k|\underline{\boldsymbol{x}}) = \frac{P(\underline{\boldsymbol{x}}|k) \cdot P(k)}{P(\underline{\boldsymbol{x}})}\,. \tag{14.9}$$

In this formula, the *a priori* probability $P(k)$ is used to include knowledge from the next level of the hierarchy. The denominator $P(\underline{\boldsymbol{x}})$ is of no interest, as it is independent of the class k. The most complicated task is estimation of the likelihood $P(\underline{\boldsymbol{x}}|k)$ or the corresponding probability density function $p(\underline{\boldsymbol{x}}|k)$. A commonly used solution assumes that the model $\mathcal{M}_\mathrm{k}$ for the class k is established by a first order Markov source followed by a transmission channel which transforms the output of the Markov source into feature vectors according to predefined statistics. The model is called the *hidden Markov model* (HMM), as the feature vectors but not the output of the Markov source can be observed. There are well-established algorithms for learning the parameters of the model from the training set and for calculating the likelihood $P(\underline{\boldsymbol{x}}|k)$ from those parameters. They are described in several books, such as those by Rabiner and Juang (1993) or Deller *et al.* (1993).

14.3.2 Principles of Speech Recognition

A speech recognition system (SRS) applies the principles of structural classification to the speech signal. It uses the components of the analysis branch in the general speech processing system in Figure 14.4, giving the structure illustrated in Figure 14.11.

The speech recognition system in Figure 14.11 is considerably more complicated than the simple pattern recognition system illustrated in Figure 14.6. However, the structure of the speech recognition system can be modified to suit the specific application. Table 14.4 summarizes the three main criteria which contribute to the complexity of a speech recognition system. The complexity of an application depends on its real requirements and their relationship to the criteria in Table 14.4. Figure 14.12 shows that some combinations of criteria are relatively easy to implement, whereas others pose significant difficulties. A number of different system classes with different degrees of implementation difficulty will be described in the following section.

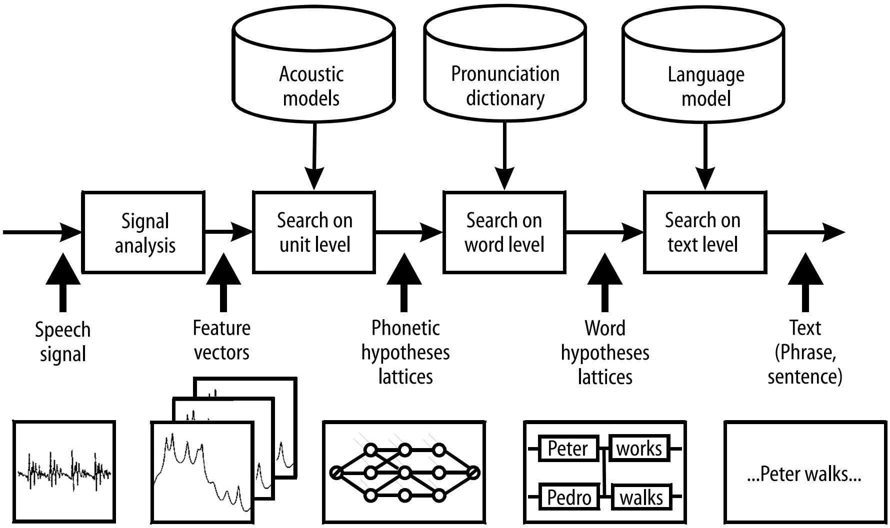


Figure 14.11. Structure of a speech recognition system. This scheme corresponds to the analysis branch of the UASR shown in Figure 14.4

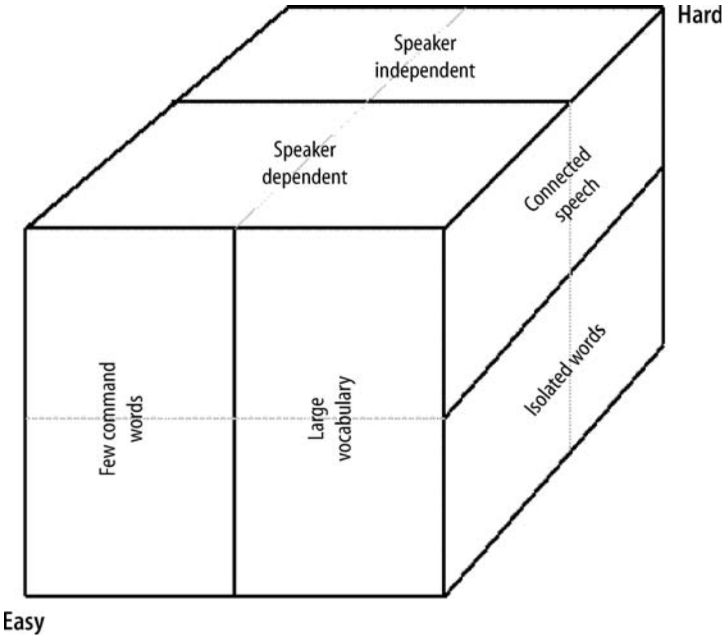


Figure 14.12. Different classes of speech recognizers resulting from the criteria of Table 14.4

Table 14.4. Criteria for the complexity of speech recognizers

Criterion	Easy	Hard
Vocabulary size	*Few number of words* Application: Command & control Simplification: Acoustical modelling at word level Training: Speaking the reference words	*Large vocabulary* Typical application: Dictation machine, number of words: more than 10,000 Problem: User cannot speak so many words in training phase Solution: Learn phoneme models, use a pronunciation dictionary
Speaker dependency	*Speaker dependent* Application: Single user systems Simplification: Only intraindividual differences have to be considered in the models Training: Learning or (at least) adaptation of the acoustical models from speech of the special user	*Speaker independent* Application: Universal Problem: Interindividual differences have to be modelled also Solution: Learn word or phoneme models, resp., using large databases from many different speakers
Naturalness of speaking	*Speaking isolated words* Simplification: Because the boundaries of the words are known, the structure of the recognizer is simplified essentially	*Connected (fluent) speech* Segmentation problem: Speakers do not leave pauses between spoken words (other than leaving blanks between words when writing) Solution: Apply the complete hierarchical structure including word hypotheses graphs and language model

14.3.3 Equipment and Applications

Command and control (C&C)

There are a number of different situations where it is not practicable to use a keyboard or other complicated setup for control. These situations can be divided into the following two main categories:

- The user is concentrating on other, more important parts of the process (industrial equipment, medical devices, components in cars).
- There are adverse environmental conditions (*e.g.* in certain industrial and/or outdoor applications).

In these “hands free” applications, speech provides a useful input method. They are characterized by a limited number of short commands, typically single words or short phrases. Therefore a recognizer with a relatively simple structure is sufficient, as the components of the scheme of Figure 14.11 are only required up to the word level (for command or word recognizer). The phoneme level and the

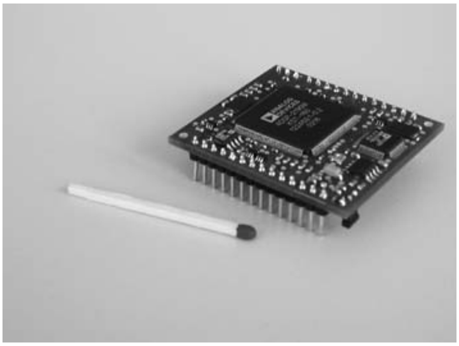


Figure 14.13. Hardware platform for a recognizer for command and control. This example shows a board consisting of a 16 bit fixpoint DSP, a 16 bit codec, and an external flash memory (mounted at the rear of the board). The board is designed to run the C&C recognizer verbKEY along with noise suppression and voice prompting software. Reproduced by permission of the voiceINTERconnect GmbH, Dresden, Germany

word level components of this structure are merged in most cases as the words are immediately used as the acoustic units in the database.

Due to this simple structure, a command recognizer is suited for applications with low computing power and/or memory resources. This situation is typical in *embedded systems* (see Figure 14.13 for an example). However, the benefits of a simple structure are partially cancelled out by the requirement for additional software components to improve the input signal. This is due to the fact that this type of speech recognizer is frequently used in acoustically adverse and noisy environments, such as cars and workshops.

Dialogue systems

Speech-based information and/or ordering systems are used in a number of other application areas of speech technology. Telephone-based systems are ideally suited for this technology as their single communication mode is speech.

The example in Figure 14.14 contains all the essential components of a dialogue system:

- The recognition branch (lower part of the illustration) consists of a recognizer for fluent speech combined with a subsystem which extracts the essential contents of the utterance. This means that the whole system configuration of Figure 14.11 is required. As only the contents of the utterance rather than the complete transcription are required, the last block of Figure 14.11 is slightly modified to perform the contents analysis. In the example in Figure 14.14, the system needs to know the following information: climate = warm, time = end-of-March, category = cheap. Sometimes it is sufficient to search for a set of relevant words. This procedure is called *word spotting*.
- The dialogue management system uses the results of the recognition branch to access the database. This is comparable to the behaviour of a human operator, who is replaced by the automatic system.

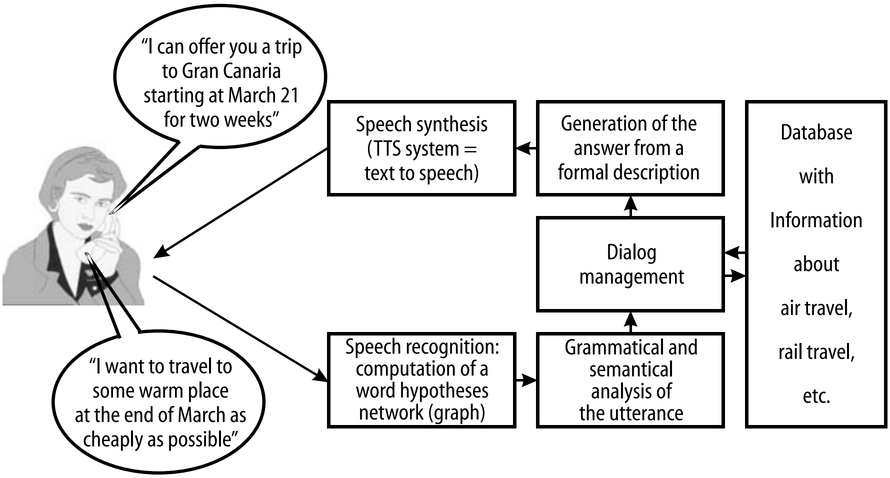


Figure 14.14. Using a speech dialogue system for ordering travel *via* telephone (adapted from Fellbaum 2000)

- The synthesis branch (upper part of the illustration) generates the answer based on the information extracted from the database. A text-to-speech system then transforms the answer into a speech signal. (The processes involved will be discussed in Section 14.4.)

It should be stressed that whether or not the system is acceptable to end-users is strongly influenced by the design of the dialogue structure.

Dictation systems

When the idea of speech recognition was first considered several decades ago, the main interest was in an automatic dictation system that could replace a secretary. However, this has proved to be a very challenging goal, as a universal dictation system would require unlimited vocabulary. A recognizer for fluent speech with very large vocabulary is located in the upper right corner of the "cube of difficulty" in Figure 14.12, associated with the most complicated case. Because this application requires all the components of the speech-to-text conversion process, the dictation system requires the full system structure of Figure 14.11. The databases are very large because they contain several thousands words.

Despite this complexity, the increasing performance specifications of personal computers, have allowed the market introduction of a number of commercial dictation systems with large vocabularies. Analysis of ten commercially available German language dictation systems (Flach *et al.* 2000) has found a word recognition rate of 80–85%. This is very impressive, considering the complexity of the task, but still too low for widespread application. However, dictation systems are gradually improving and improved user interfaces facilitate the identification and correction of speech recognition errors.

Some developers of dictation systems have improved product performance by developing domain specific and therefore reduced vocabularies. There are several successful systems of this type, for instance in medical applications for the dictation of radiological or other medical results and in lawyers' offices.

Speech understanding and translation systems

A number of science fiction films illustrate the final goal of speech technology of the user having a conversational interaction with the computer. A classical example is HAL 9000 in Stanley Kubrick and Arthur C. Clarke's *2001: A Space Odyssey*. The year 2001 is now in the past, but HAL has not yet appeared. It should be noted that this task requires the conversion of speech-to-text followed by a subsystem for *understanding* the text.

The block diagram in Figure 14.11 is not sufficient for a speech understanding system. It must be complemented by blocks mimicking the higher level processes of human speech perception. The degree of complexity of these components is very variable and depends on the application. A simple example of a rudimentary comprehension component, which is very far from the requirements of HAL, is the dialogue component of Figure 14.14. It should be noted that the field of speech understanding is part of research in artificial intelligence (AI).

Automatic translation is one of the cases where contents analysis is required. Automatic translation in AI research has traditionally focused on text-to-text translation. The extension to speech-to-speech translation poses a number of complex additional problems, as the input to the translation component of the system may not be grammatically correct, due to a combination of spontaneous speech and speech recognition errors. Several large research projects have produced results on the potential and limitations of speech-to-speech translation. For instance, the German project Verbmobil (Wahlster 2000) investigated bidirectional translation tasks with spontaneous speech input in negotiation dialogue in a restricted domain with a vocabulary size of 10,000 word forms. The system was evaluated by counting the number of "approximately correct" translations, defined as translations which convey the speaker's intended effect to the recipient. Verbmobil achieved more than 80% approximately correct translations in its restricted domain.

It should be noted that speech-to-speech translation is one of the most challenging, as well as one of the most promising research areas with useful applications in the field of speech technology. If they become widely available, automatic translation systems are likely to have significant impacts on society worldwide, discussion of which is beyond the scope of this book.

14.4 Text-to-speech Conversion

14.4.1 Principles of Speech Production

Human and synthetic speech production

Human speech production is a very complex process (Levelt 1989). The complex steps required to produce an utterance can be divided into the following two categories:

- The planning and decision processes in the brain required to produce a formulation following the grammatical rules of the relevant language from semantic contents or an intention to speak.
- Activation of the muscles controlling the breath and the synchronous movement of the articulators to produce an acoustical waveform which is radiated by the mouth.

There is an area of AI called *generation* which models this complex interaction of thinking and speaking. Its main aim is the conversion of nonverbal information into natural language (Görz *et al.* 2000). An illustrative system which includes a generation component has already been discussed and illustrated in Figure 14.14. Coupling the generation component with a speech synthesizer produces a *contents-to-speech* or *concept-to-speech* system (CTS system) which models the process of human speech production. However, in many applications, the input information is already available in the form of written language (text) and only the simpler structure of a *text-to-speech* (TTS) system is required.

Text-to-speech systems

This section will discuss the main principles of TTS systems. The block diagram of a TTS system is derived from the right (synthesis) branch of the general speech processing system in Figure 14.4, giving the structure in Figure 14.15.

Comparison of Figures 14.4 (universal analysis and synthesis system), 14.11 (speech-to-text) and 14.15 (text-to-speech) shows that the inclusion of prosodic information (intonation, sound duration, sound energy) is expressed by a separate box in Figure 14.15. Careful and correct treatment of prosodic elements increases the naturalness of synthesised speech and this is important for user acceptance. As discussed in Section 14.2.3, this task is not easy and a body of research in the last decade has focused on improving the quality of prosody in TTS systems (for example, see Hoffmann *et al.* 1999a).

The most crucial part of a TTS system is the rightmost box in Figure 14.15 which aims to produce a speech signal (*acoustical synthesis*). There are two approaches, parametric and concatenative speech synthesis, which will be discussed in the next section.

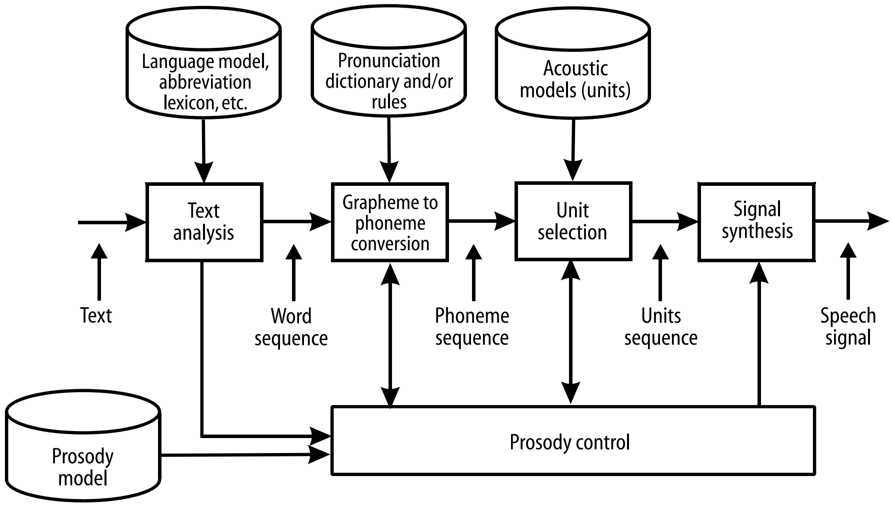


Figure 14.15. Structure of a TTS system. This scheme corresponds to the synthesis branch of the UASR shown in Figure 14.4

14.4.2 Principles of Acoustical Synthesis

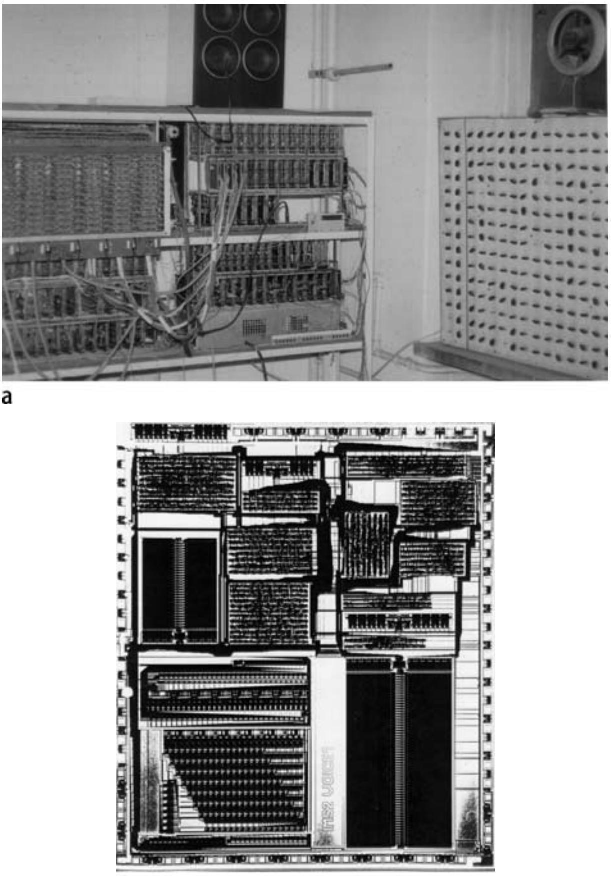
Parametric speech synthesis

The construction of a technological system to produce speech can be performed in several ways. The most obvious solution applies a model of the human articulation system. It is necessary to control the parameters of this model in order to produce the different speech sounds. This concept of a *parametric speech synthesizer* was discussed briefly in Section 14.2.1 with reference to Figure 14.3c. In this special case, it was necessary to control the following parameters: the filter parameters, the gain, the position of the switch between voiced and unvoiced sounds and the frequency of the generator for voiced sounds.

The idea of parametric speech synthesis considerably predates electronics and the most successful and famous mechanical model of the articulation system was invented by Wolfgang von Kempelen (Kempelen 1791).

Since the problems of storage and transmission of speech were satisfactorily resolved relatively early in communications engineering, there were several early attempts at electronic speech synthesis. The first electronic speech synthesis systems were a consequence of the development of powerful transmission systems based on a German patent (Schmidt 1932). The earliest implementation was Dudley's Vocoder (voice coder) in 1936.

Following the development of the Vocoder, a number of parametric synthesis systems along the lines of Figure 14.3c were produced. The same principle can be found on all the hardware platforms produced in recent decades from electronic valves *via* discrete transistor circuitry, integrated circuits and microprocessors to state-of-the-art DSPs (Digital Signal Processors). Figure 14.16 illustrates this



a

b

Figure 14.16a,b. Selected formant synthesizers developed at the TU Dresden. **a** partial view at the three formant synthesizer SYNI 2 from the year 1975. This device was basing on germanium transistor technology and was controlled manually or by a paper tape reader. The following synthesizer versions were computer controlled according to the availability of process-control computers or, later on, microprocessors. **b** this layout photo shows the final point of this line, the formant synthesizer chip VOICE 1 which was developed with the Fraunhofer Institute IMS in Dresden

development by means of an example. In this example, the linear filter of the block diagram in Figure 14.3c is designed to produce a formant structure based on Figure 14.7 for *formant synthesis*. It should be noted that only two decades of development separate the very different implementations shown in the two photographs.

The development of computers made possible the use of parametric speech synthesis, as they could be used to send the control parameters to the synthesizer hardware in real-time. However, the quality of the synthesized speech was poor due to significant differences between the human speech production system and the model.

Concatenative speech synthesis

The limited quality of parametric speech synthesis led to repeated attempts to synthesize speech by concatenating short segments of speech previously spoken by real speakers. Before the development of digital computers this so-called *synthesis in the time domain* required very complicated analog equipment. The introduction of computer control did not resolve the problems immediately due to the limited magnetic core memory of only a few kilobytes of the early process-control computers, such as the PDP-8 series. This is clearly insufficient for storing even a very short digitized speech signal, since, as shown in Section 14.2.1, 1.3 s of speech requires more than 20,000 samples and approximately 40 kB of memory. Therefore, the broad development of synthesis in time domain or *concatenative synthesis* started later when cheap semiconductor memories were introduced resulting in the development of personal computers.

Producing concatenated speech from the stored waveforms of single sounds (allophones) poses a specific problem. Real speech is highly dynamic, and the transitions from one sound to the next (the effects of *coarticulation*) are difficult to model in the concatenation software. As a standard solution to this problem, combinations of two sounds, called *diphones*, are selected from natural speech and stored to form a *diphone inventory* for the synthesis system. The waveform of a diphone starts in the middle of the first sound and ends in the middle of the second sound, as shown in Figure 14.17.

A complete utterance can now be produced by forming a series of the corresponding diphones. To minimize audible distortions, the diphones cannot be linked together without a certain overlap. This is performed by an *overlap-and-add* algorithm (OLA). The most commonly used OLA algorithm is known as TD-PSOLA (time domain period synchronous OLA). Its description can be found in textbooks such as (Dutoit 1997). This method gives a smooth concatenation and also allows the duration of the signal to be lengthened or shortened to a certain extent. This is required to control the fundamental frequency (pitch) which is the most important prosody parameter (see Section 14.2.3).

There is a potential danger of reduction in quality of the synthesized speech at the concatenation points. Other concatenation algorithms than PSOLA have been

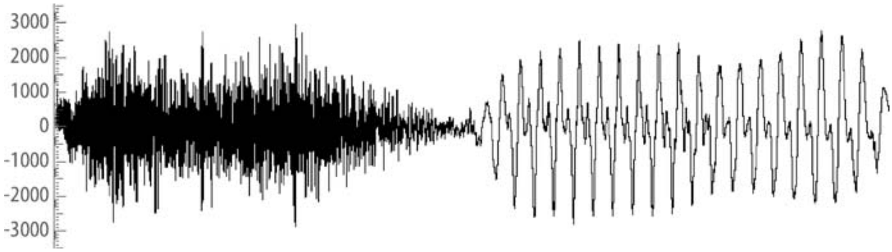


Figure 14.17. Example for a diphone. It shows the right half of the fricative sound [f] and the left half of the vowel [u] from the English word *beautiful* which was pronounced by a female native speaker. The diphone length is approximately 200 ms

Table 14.5. Inventories of the TTS system of the Dresden University (Dresden Speech Synthesis DRESS) as available in 2003

Language	Number and type of speech units	Speaker	Size (16 Bit PCM)
German	1212 diphones	1 male, 3 female	5 MB
US English	1595 diphones	1 female	7 MB
Russian	572 allophones	1 male	0.5 MB
Mandarin Chinese	3049 syllables	1 male	27 MB
Italian	1224 diphones	1 male	4 MB
Klingon	299 allophones	1 male	1.3 MB

implemented with varying degrees of success. Speech quality can be improved by avoiding concatenation points by using speech units which are longer than diphones. However this increases the number of units and the memory space required for their storage. The complete set of units (whether diphones or other speech units) is called the *inventory*. It forms the acoustical database in the block diagram in Figure 14.15. To give an idea of the order of magnitude of the memory required, Table 14.5 characterizes the inventories for a particular TTS system (Hoffmann *et al.* 1999a). It should be noted that the table includes databases of very different types of units and consequently of very different memory sizes.

14.4.3 Equipment and Applications

Performance criteria

The structure of state-of-the-art TTS systems follows that of Figure 14.15. In nearly all cases, the acoustic component is implemented using the concatenative method in the time domain. The following criteria can be used to evaluate the performance of a TTS system:

- *Intelligibility*. This is the primary criterion but it only plays a minor role in contemporary discussions as most TTS systems produce speech with good intelligibility.
- *Naturalness*. Listeners are very sensitive to the phenomena which make synthetic speech sound unlike that of a natural speaker. Naturalness is influenced by both the segmental quality (the quality of the units forming the acoustic database) and the prosody model. It is also very important that the text analysis block of Figure 14.15 supplies the prosody control module with exact input information. Increased naturalness can also be achieved by modelling the so-called *spontaneous* effects in human speech. They mainly include speech rhythm and variations in pronunciation (Werner *et al.* 2004).
- *Multilinguality*. TTS systems which can produce speech in different languages are very useful and of greater commercial value. This requires one set of databases in the scheme of Figure 14.15 for each language (cf. the example

in Table 14.5). Changing the language simply requires the system to change databases. However database memory requirements for this type of system increase linearly with the number of languages and therefore a "universal" inventory containing the sounds of all the languages required for a particular TTS application has been investigated. The quality of this so-called *polyglot* synthesis was very limited.

- *Speaker characteristics.* As shown in Table 14.3, the speech signal carries para- and nonlinguistic information. Therefore, there are advantages in providing TTS systems with different speaking styles and the ability to express emotions and offer different speakers. Again, the simplest way to do so is to implement different databases which can be drawn on for the specific configuration. However, this solution leads to the same memory problem as the introduction of different languages and therefore researchers are investigating algorithmic methods for influencing the signal directly. For instance, it is possible to switch from a male voice to a female one and *vice versa* using an algorithm which takes into account the gender specific differences of the human vocal tract length (Eichner *et al.* 2004).
- *Resources.* The cost of a TTS application depends on the quantity of resources required by the system. These resources include the computing power required to calculate the synthetic speech signal in real-time and the memory required by both the algorithms and the databases. It is frequently the memory requirements which have the greatest impacts on the total cost. The total amount of memory required by a TTS system is called its *footprint*. The cost of the system is particularly important in bulk applications like TTS in mobile phones or other embedded solutions.

Diphone-based TTS systems became established as the baseline technology during the 1990s. Current developments are largely influenced by the need to reduce resource requirements. There are currently two different classes of TTS systems, PC- or server-based systems and embedded systems, with different types of trade-offs between speech quality and resource requirements in the two cases.

PC- or server-based systems

Resource requirements are not particular significant for TTS systems which run on a PC or workstation. Therefore it is feasible to use speech units which are larger than diphones in PC- or server-based TTS systems. In extreme cases, very large databases (*corpora*) of recorded speech are used. For instance, the Verbmobil system (Wahlster 2000) used a *corpus* of 3 h of speech to synthesize utterances from its restricted domain. However, such *corpora* require gigabytes of memory to achieve their aim of producing very natural sounding speech and still have a number of unsolved problems (Hess 2003):

- *Cost functions.* A given text can be synthesized by different segments of the *corpus*. The selection of the non-uniform segments that are best suited to the given utterance is a complicated optimization problem, which can be solved using cost functions. There is still ongoing research on the design of these cost functions.

- *Coverage*. Even large *corpora* cannot contain all the word forms of a given language. Languages have a large number of rarely used words. If these rare words occur in a text to be synthesized, they must be composed of smaller units such as syllables or diphones. This results in the concentration of a large number of concatenation points in a specific part of the synthesized speech and the danger of audible effects.
- *Labelling*. Before speech segments from the *corpus* can be used, they must be identified and labelled accordingly. This cannot be carried out manually for large *corpora*, but the quality of automatic labellers is still unsatisfactory.

Embedded systems

In the other cases of TTS systems which are integrated (embedded) into PC-independent low-cost applications the costs of the system are largely determined by its footprint, which should therefore be minimized as far as possible. As an illustration, Table 14.6 gives the costs of on-chip RAM for several different memory capacities.

A complete TTS system with a footprint of less than 1 MB was achieved for the first time with the microDRESS version of the TTS system DRESS (presented in Table 14.5) (Hoffmann *et al.* 2003). Since 1 MB is rather small compared with the memory requirements of the data in the baseline system in Table 14.5, it is necessary to concentrate effort on reducing the inventory size. This can be done in two steps. The first step generally involves a reduction to telephone quality covering the bandwidth from 300 Hz to 3400 Hz, with an associated reduction in the number of samples required for the inventory. The second step involves the application of coding algorithms, which are frequently used in communication systems. The application of simple coding schemes is sufficient to give a footprint of 1 MB. A smaller footprint could be obtained by the use of more powerful coding algorithms, but would have the disadvantage of reducing the speech quality as is known from the speech encoders and decoders (combined known as *codecs*) which are applied in telecommunications (Chu 2003). This reduction in quality can be offset to a limited extent by carefully maintaining the inventory.

Table 14.6. On-chip RAM as cost driver. Data from the year 2002 (Schnell *et al.* 2002)

RAM (kB)	0	120	144	292	512	1024
Costs (EUR)	0.83	1.27	1.49	1.93	3.19	5.00

TTS systems compared to recorded speech

The availability of large capacity cheap memory chips has made it possible to store whole spoken utterances, thereby eliminating the need for TTS synthesis under the following conditions:

- The speech *corpus* consists of a fixed set of utterances.
- The speech *corpus* is sufficiently small to have limited memory requirements for storage and limited time requirements for production by a human speaker.
- The *corpus* is fixed, with low probability that it will be necessary to change it in the future.

Recorded speech has a number of applications in daily life, such as public announcements at stations. The principle is known from voice recorders for storing short dictations or other audio information. These recorders are available as both small stand-alone devices and integrated into pocket PCs and other applications. Recorded speech can be of hi-fi quality or reduced quality due to degradation by the codec used to compress the signal. However, recorded speech always sounds more natural than synthetic speech. This can lead to problems in evaluating the quality of TTS systems if users are not aware of the difference between recorded and synthetic speech.

14.5 Braille Conversion

14.5.1 Introduction

Braille is used to represent text by means of tactile symbols. Since tactile text reading is discussed in Chapter 4, it will only be considered briefly here. The main idea was developed (from an unsuccessful system developed for the French army and called night writing) by the French blind teacher Louis Braille in about 1825 and aimed to replace the visible characters in text by a tactile array of up to six dots. Braille is still an important means for blind people to access written documents, though the proportion of blind people who read Braille is low. The elements of a printed text are called characters or signs, whereas the elements of a Braille transcript are called symbols.

Braille code is physically implemented in the form of small raised bumps on paper or other material, which are touched with the fingertips. The bumps can be produced using a stylus and a slate. The slate comprises an upper part with rectangular openings to ensure appropriate positioning of the symbols, and a lower part with small indentions for positioning the dots. Producing large texts in this way is both time-consuming and complicated, as a mirror writing approach must be used for punching the Braille symbols to produce raised bumps rather than depressions. Therefore, mechanical writers called Braille writers or Braillers are used. They are similar to a typewriter with six keys (one key for each dot of a Braille symbol) and a space bar, as illustrated in Figure 14.18. They are operated by Braille transcribers,

a b

Figure 14.18a,b. Old and new Braille writing devices: **a** mechanical writer, embossing on a paper tape for stenographic purposes (Firma Karl Neubert, Leipzig, around 1950), from the historic collection of the TU Dresden (photograph by R. Dietzel); **b** electronic Braille writer Elotype 4E, also applicable as computer-controlled Braille embosser (by courtesy of Blista Brailletec, Marburg)

who are usually certified professionals, as they have to perform the transcription process mentally, which is not carried out by an automatic (technical) system.

Automatic systems for transcription from text into Braille and Braille into text have a number of advantages, particularly with regards to speeding up the process, but require suitable input and output devices. Other than the need for appropriate hardware, the conversion seems relatively simple, as it is performed sign by sign at the text level. As will be pointed out in the following, this is not in fact totally the case.

The original approach was based on a one-to-one coding of the characters with code combinations defined for all letters, numbers, punctuation marks, and a few special symbols, such as those indicating capitalization. This baseline system, which includes a few elementary composite signs, is called Grade 1 Braille. Producing text in Grade 1 Braille requires a lot of space. One page of a Braille book contains approximately 1000 Braille symbols, whereas one page of a printed book may contain 3500 characters. Therefore a number of rules were introduced to reduce the volume of a Braille coded text by 20–30%. (This is comparable to using shortcuts such as “asap” for “as soon as possible” or the sign “4” for the word “for” in SMS or e-mails.) This has resulted in an extensive system of so-called contractions and the rules for their application. For instance, if “&” can be used for “and”, then “h&le” stands for “handle”.

The complete system is called Literary Braille Code, Grade 2 Braille or contracted Braille. Most blind people who read Braille can also read the contracted version. Unfortunately, the contractions are language specific and therefore the code for contracted Braille must be learnt for each language separately. In German speaking countries, the terms *Vollschrift* (full text) and *Kurzschrift* (short text) are used for Grade 1 and Grade 2 Braille. Different organizations coordinate the standardization in the different countries, as shown in Table 14.7. Different approaches are required in some non-European language systems. For instance, Chinese Braille symbols represent the sounds of the spoken language rather than the characters of the written language.

Table 14.7. Examples for organizations and resources of the Braille system for different languages

Language	Organization	Examples for web resources[a]
English	International Council on English Braille (ICEB) Braille Authority of North America (BANA) Braille Authority of the United Kingdom (BAUK)	www.iceb.org www.brailleauthority.org/ www.brl.org/ebae/ www.bauk.org.uk/
French	Fédération des Aveugles et Handicapés Visuels de France	www.faf.asso.fr/sommaire.htm
German	Brailleschriftkommission der deutschsprachigen Länder	www.fernuni-hagen.de/ZFE/fs/download.htm
Chinese	China Association of the Blind	www.braille.ch/pschin-e.htm www.omniglot.com/writing/braille_chinese.htm www.hadley-school.org/Web_SIte/8_d_chinese_braille_alphabet.asp
Japanese	Japan Kanji-Braille Association	kantenji.jp www.geocities.co.jp/CollegeLife-Library/7524/tenji/tenyaku.html

[a] A general overview on the codes in various languages is presented in homepages.cwi.nl/~dik/english/codes/braille.html

There are also special forms for the following purposes, amongst others:

- Musical notation (www.brl.org/music/index.html; for details see Chapter 16).
- Mathematical notation; although there is a universal mathematical text notation, this is unfortunately not the case for Braille and different systems of mathematical Braille have developed in different countries. This has a number of disadvantages, including making both the production of mathematical Braille and communication between blind mathematicians in different countries more difficult. The US version is known as The Nemeth Braille Code.
- Scientific notation, for instance for chemistry (www.brl.org/chemistry/).
- Phonetic notation (www.clauchau.free.fr/L/phonalph.html).
- Computer Braille, which uses an 8 dot system for 256 signs, analogously to the ASCII code (see www.braille.org/papers/unive/unive.html or DIN 32 982).

With regards to text-to-Braille conversion, there are clearly additional difficulties associated with the production of contracted Braille. Some remarks on this topic can be found in the first volume of this series (Hersh and Johnson 2003, pp 262–265). The following two subsections consider hardware and software aspects of text-to-Braille and Braille-to-text conversion. It is often useful to provide both Braille and speech output and examples of dual output systems are presented in Section 14.6.

14.5.2 Text-to-Braille Conversion

Historical development

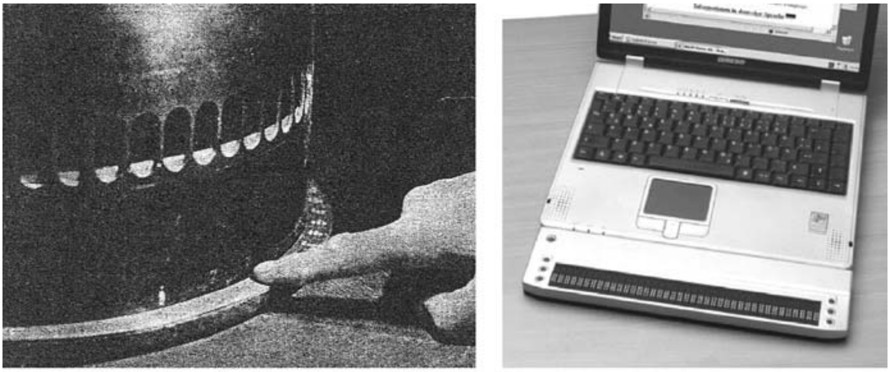
Analogously to speech-related technologies, the development of modern reading machines has required the availability of computer technology. However, the requirement for tools for blind people to read printed texts and thus enhance their independence was recognized decades earlier. The discovery of the photoelectric effect allowed printed signs to be converted into electrical signals, which were presented in audio form in the earliest reading devices. Early devices with audio output included the following:

- The *Optophone* was invented in 1914 by E.E. Fournier-D'Albe, a Professor in Birmingham, UK. In the optophone the printed text characters are irradiated with light pulses and the reflected light is converted into an audible signal by means of a Selenium cell. The listener uses the different sounds to distinguish between the different text symbols. The invention was subsequently modified and applied in several different devices.
  (See www.oldweb.northampton.ac.uk/aps/eng/research/optophone/optophone2.html.)
- Another early reading machine was constructed by Rosing, a Professor in Petersburg before 1917. The text characters were scanned by an optical beam in the horizontal and vertical directions to produce a Morse-like sound.

The complexity and lack of an intuitive relationship between the text and the acoustic patterns and the lack of synthetic speech output led to the development of reading devices with a tactile output, including the following:

- W. Thorner, an eye specialist in Berlin, obtained a patent in 1916 for converting images to electro-tactile patterns. Later inventions used grids of electromagnets for the production of mechanical stimuli. This approach is still used for presenting graphical information (see Chapter 4 of this book). However, it proved to be unsuitable for reading text unless the magnetic elements were combined to give Braille symbols.
- G. Schutkowski, a teacher at a school for blind people near Berlin, was probably the first person to recognize that the most appropriate tactile output for reading machines was Braille. He invented an optical pattern matching method for reading (Schutkowski 1938). The reading mechanism was coupled with a drum which moved the corresponding Braille symbol in front of the user (Figure 14.19a). Remarkably, he also investigated the addition of speech output to his device (Schutkowski 1952) and realized that the spelling mode which he developed was only the first step towards a full text-to-speech system.

The current situation is summarized in Figure 14.20. Technical support systems for a human transcriber are presented in the upper part of the figure. The Brailler has already been mentioned. It can be replaced by a PC with driver software which changes the standard keyboard to a six-key input device. Automatic text-to-Braille conversion is illustrated in the lower part of the figure. The components of this process will now be described.



a b

Figure 14.19a,b. Old and new devices for volatile Braille output: **a** the drum of the early prototype of a text-to-Braille machine by G. Schutkowski (contemporary press photograph from the early 1950s); **b** a Braille display with 40 elements (type SuperVario), mounted in front of a PC keyboard (by courtesy of BAUM Retec AG)

Input devices

The text to be converted needs to be available in machine-readable form. If this is not the case, a text file can be produced in one of the following standard ways:

- By using a PC with a conventional text processing system.
- By using a scanner to input the printed text. The scanner is connected to a standard PC which hosts the optical character recognition (OCR) software which produces the text file. Since errors cannot be completely avoided with the current state of OCR technology, the word processing software of the PC should be used to edit the resulting file and correct any errors.

The scanner is the essential input device for a reading machine. Since it performs similarly to other OCR applications, its properties will not be discussed here. They are described in more detail in Chapter 15 and the reader is referred to Bunke and Wang (1997) for additional information.

Output devices

The output device used depends on whether a durable but bulky paper version of the converted text or a quick to produce but transient (volatile) version is required.

Producing Braille symbols on paper using an automatic system requires the manual keys of the traditional Brailler to be replaced by a corresponding computer interface, called a *Braille embosser*. A number of different types of embossers are available and an illustration is presented in the first volume of this series (Hersh and Johnson 2003, p 265). See also Chapter 12 of this book. Some types can produce dots on both sides of the page (interpoint Braille). An electronic Brailler can be used as a computer output device also (Figure 14.18b).

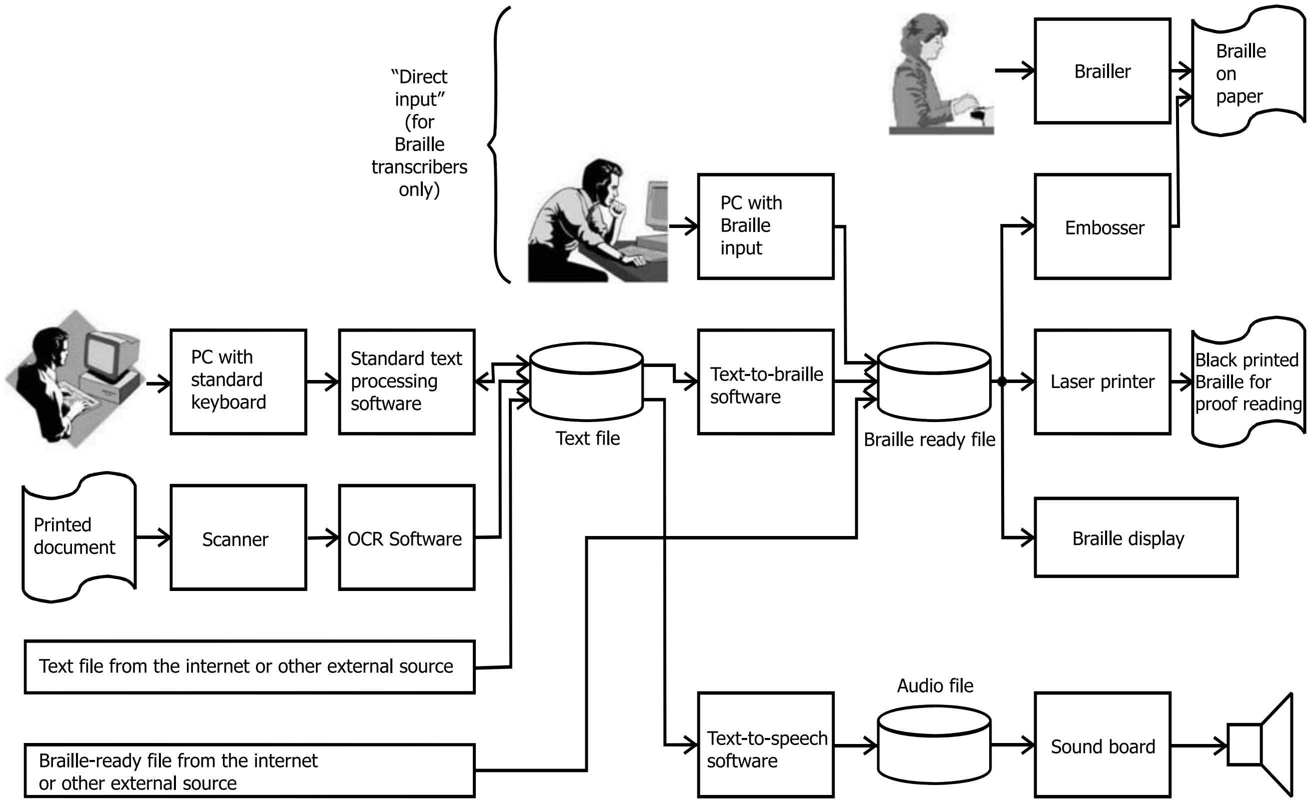


Figure 14.20. Data flow in text-to-Braille processing. The diagram shows how the work of human Braille transcribers (*above*) is automated by suited hardware and software components. Additionally, the inclusion of text-to-speech components in the system is illustrated

Braille displays are used to provide the user with refreshable (volatile) Braille output. They consist of cells which each produce a touchable Braille symbol by using a piezoelectric drive to raise a number of small pins. Displays consist of lines of 20, 40, or 80 cells, as shown in Figure 14.19b. Braille displays were introduced by the company F.H. Papenmeyer in 1975. Modern Braille cells have 8 dots and can therefore be used for computer as well as literary Braille.

Conversion software

Text-to-Braille conversion software is the main component of Figure 14.20. It needs to carry out the following different groups of tasks:

- *Mapping the text characters to Braille symbols*. This is the central module of the conversion process. It should be noted that this mapping is not one-to-one mapping to the use of a large number of contractions and special rules, which are language-dependent as it was explained already.
- *Text analysis*. The use of contractions is governed by a number of specific rules. These include shortcuts for prefixes and suffixes, preference rules for ambiguous sequences and contractions which cannot be applied over syllable boundaries. Therefore, the text-to-Braille mapping module needs to "know" the morphological structure of the words and this is the task of a special module. Morphological analysis is a typical AI problem, which is apparently relatively easy to carry out for a person, but difficult for an automatic system. The accuracy of the text-to-Braille conversion depends on the morphological analysis. However, there are frequently morphological ambiguities, which can only be resolved only at the semantic level. Morphological analysis is also a component of the text analysis module of a TTS system, as shown in Figure 14.15. In both cases, certain types of errors are difficult to avoid and lead to prosodic errors in TTS and transcription errors in text-to-Braille conversion.
- *Producing the correct layout from the Braille symbols*. There are clearly defined layout rules for books and other documents in Braille in order to facilitate navigation and structure recognition by blind readers. The conversion software must also consider these layout rules, particularly if an embosser is used to produce a paper document. The layout rules for English Braille documents are summarized in the document "Braille Formats: Principles of Print to Braille Transcription 1997" (www.brl.org/formats/).

As shown in Figure 14.20, the conversion software works on a file-to-file base. This has two main advantages and one drawback:

- Using a text file for input is convenient for storing and re-using the data and is also required to enable removal of any errors introduced by the scanner OCR software using word processing software. This is easier than correcting the Braille output.
- The Braille file can be evaluated and corrected using special software tools. This is necessary to remove errors from the text-to-Braille conversion. These

tools can also be used to print the Braille output in black on a laser printer for proofreading by sighted users.

- The drawback is due to slight incompatibilities between the computer code (ASCII) and the Braille code which leads to additional ambiguities. For example, the single ending quotation mark and the apostrophe are the same in ASCII, but different in Braille (Gray 1998).

Figure 14.20 does not include conversion of the contents of a computer screen or web documents. Reading the screen contents is the main application of Braille displays. As this chapter is concerned with speech and language, the related issues of interpreting layouts, embedded graphics and formulae will not be discussed and the reader is referred to Chapter 13. A discussion of the different formats which can be used to represent the input of the conversion process (ASCII, MS Word, LaTeX, HTML, and SGML) have been given by Kahlisch (1998).

There are differences in the software required according to whether the Braille output is directed to an embosser or to a Braille display. These details are not considered in Figure 14.20. When combined with a computer keyboard, a Braille display should be able to switch between the following modes:

- Navigation (exploring the contents of the screen).
- Communication with a program.
- Reading long texts.

14.5.3 Braille-to-text Conversion

The (re)conversion of Braille to text is required in a number of applications. The main application is communication by public organizations with blind people. Other applications are in education and research, where sighted people may want to read Braille documents. Automated Braille-to-text conversion does not require any special hardware, as a standard flat-bed scanner can be used to input the Braille symbols and further processing carried out on a PC. The standard OCR software for the scanner needs to be replaced by software that is able to recognize Braille symbols. After recognition by the OCR software, the sequence of Braille symbols is transformed to the corresponding sequence of text characters. The Braille recognition and transformation software will be described in more detail below.

Another application of Braille-to-text conversion, which only requires the second (transformation) step, is the conversion of Braille files into text files for subsequent text processing or other electronic applications. Examples include production of the minutes of a meeting using a portable electronic Braille notetaker. The Braille minutes are transferred from notetaker to a PC which produces a Braille text file which must be converted to a text file before it can be distributed to sighted partners by e-mail or in paper format.

OCR software

Other than in text reading, Braille symbols embossed on a thick sheet of paper cannot be distinguished from the background by their colour and the Braille dots must be identified by their appearance in the light of the scanner lamp. Many Braille documents are in "interpoint Braille", with Braille symbols embossed from both sides of the paper to utilize the voluminous carrier material more effectively. This requires the OCR software to distinguish between the raised bumps and depressions. This poses additional challenges compared to the text-based OCR presented in Chapter 15. The essential steps in the recognition process include the following (Antonacopoulos 1997):

- *Image acquisition*. The camera obtains a grey-level image with each dot represented by a light area combined with a shadow. The details of the image will depend on the type of scanner and its angle of illumination. A spatial resolution of between 80 and 200 dots per inch is recommended.
- *Pre-processing*. The image quality is improved using image processing methods.
- *Identification of the dot*. The software identifies the regions of the grey-scale image which correspond to a dot.
- *Segmentation of the Braille symbols*. The detected dots are grouped as Braille symbols or *cells*. The decision on which points belong to the same cell depends on the vertical and horizontal spacing between dots. This is not an easy task as can be seen, for instance, by considering the horizontal spacing. The following distances a, b, and c are of increasing size and defined as the distance between the two columns of one symbol, a, distance between two adjacent symbols, b, and the distance between two adjacent words, c. If a symbol with dots only in its left column is followed by a symbol with dots only in its right column, the distance is $(2 * a + b)$ and this distance most be interpreted appropriately.
- *Recognition*. The positions of the symbols (cells) are known from the previous stage. The cells must be assigned a particular Braille symbol according to the presence or absence of dots in the six possible positions of the cell.
- *Verification*. Recognition errors can result from a number of different factors. Braille code does not have any redundancy at symbol level, as all possible combinations of dots are allowed. Therefore consideration of the context is required to verify the results of Braille recognition. This task can be carried out by the conversion module.

It should be noted that Braille-to-text conversion devices achieve surprisingly high recognition performance for good quality documents, whereas problems may occur for older or dirty Braille documents.

Conversion software

To conclude this section, two tasks of the module following the OCR software are discussed briefly:

- The software must perform the transformation process on a string-to-string basis in order to resolve appropriate contractions. The module for text-to-Braille conversion described above needs to consider all possible contractions and the complex set of rules for their application. The module for Braille-to-text conversion involves the inverse system.
- The software should try to recognize and, as far as is feasible, eliminate recognition errors using context analysis. Failure to eliminate recognition errors can be serious. For instance, incorrect interpretation of the special symbol that announces numbers will lead to interpretation of the subsequent symbols as letters rather than numbers. When a recognition error is suspected, the proposals for correcting it must be ranked. This is a complex task and involves consideration of the large number of possible contractions.

14.6 Commercial Equipment and Applications

14.6.1 Speech *vs* Braille

In this section, commercially available speech-related equipment aimed particularly at blind and visual impaired users will be considered. Since a fairly large number of different products are available, only a small number of selected devices will be discussed. They are divided into product groups which will be treated in the different subsections of this section. A number of application issues are discussed in other chapters of this book.

There is a degree of competition between Braille displays and speech output. Many users who are fluent Braille readers prefer to use a Braille display. However, speech output also has its advantages and can give users who do not read Braille access to computer and communication technology, including the Internet. In the U.S., for instance, according to the National Health Interview Survey, less than 10% of legally blind people can read Braille (Roe 2001, p. 109). However, it should be noted that Braille can be used by deafblind people, whereas speech output is often unsuitable for them.

Braille displays are unfortunately very expensive, with a 40-cell Braille display costing more than 5000 € and an 80-cell version about 12,000 €. However, costs could fall if technologies other than the piezoelectric effect are used in future displays. Possible technologies include memory metals and electro-rheological fluids (Weber 2004). However, speech output is currently much cheaper than a Braille display.

The speech components of assistive systems are generally adapted from general-purpose systems. A number of available speech input and output solutions are presented in Tables 14.8 and 14.9 respectively. The selection of these particular solutions was largely determined by the availability of information on the Web. The systems are representative examples of what is available, but no special endorsement or discouragement to use them is intended.

Table 14.8. Selected speech recognition systems

Name	Features	Source
Verbmobil	Translation system for spontaneous speech, using different top research systems for recognition	DFKI verbmobil.dfki.de
IBM ViaVoice	Dictation system	ScanSoft Inc. www.scansoft.com/viavoice/
VoicePro	Based on IBM ViaVoice, best dictation system according to (Stiftung Warentest 2004)	Linguatec www.linguatec.de
Dragon Naturally Speaking	Dictation system	ScanSoft Inc. www.scansoft.com/naturallyspeaking/ www.talktoyourcomputer.com/
SpeechMagic	Document creation in medical, legal and insurance solutions	Philips Speech Processing www.speechrecognition.philips.com/
StarRec	Robust speech processing mainly for automotive applications	Temic Speech Dialog Systems www.temic-sds.com/
Nuance	Speech recognition mainly for telecommunication applications	Nuance Communication Inc. www.nuance.com
VoiceMan	Voice portals for call centers and other telecommunication applications	Sikom Software www.sikom.de
SpeaKING	Speech recognition for dictation, data acquisition, control	MediaInterface www.mediainterface.de
Microsoft Speech Server (MSS)	Enables enterprises to extend Web applications to be accessible by speech	Microsoft www.microsoft.com/speech/
verbKEY	Module for robust C&C recognition, see Figure 14.13	voiceINTERconnect www.voiceinterconnect.de
VOICO	Module for robust C&C recognition	ABS Jena www.abs-jena.de

14.6.2 Speech Output in Devices for Daily Life

The following devices are used in the home environment and/or for employment. In most cases they are based on standard devices which have been upgraded by adding a speech output module. The speech output is normally recorded speech rather than TTS, as only a fixed vocabulary is required and the number of different utterances is low. Braille output is unsuitable for this type of application due to the cost of Braille displays and environmental conditions which are too rough for sensitive Braille displays. Therefore other solutions will be required for deafblind people and/or for use in noisy environments.

Examples of the use of home and employment applications of speech output include the following:

- Information on the status of a range of household devices, from a liquid level measure for use in cups and other containers to washing machines and other appliances.

Table 14.9. Selected TTS systems

Name	Features	Source
Natural Voices	Large, non-uniform speech databases for single computer applications	AT&T Labs www.naturalvoices.att.com
CHATR	Platform for speech synthesizer research, *corpus*-based	ATR Interpreting Telecommunications Research Laboratories feast.his.atr.co.jp/chatr/chatr/index.html
Festival	Free software speech synthesis workbench offering black box text to speech, as well as an open architecture for research in speech synthesis	University of Edinburgh, CSTR festvox.org/festival www.cstr.ed.ac.uk/projects/festival/
rVoice	TTS with non-uniform units	Rhetorical Systems Ltd., Edinburgh www.rhetorical.com
BOSS	"Bonn Open Synthesis System", free open source speech synthesis software with non-uniform unit selection	Universität Bonn www.ikp.uni-bonn.de/dt/forsch/phonetik/boss/index.html
MBROLA	Project for obtaining speech synthesizers for as many languages as possible, and providing them free for non-commercial applications (no text processing)	TCTS Lab, Faculté Polytechnique de Mons, Belgium tcts.fpms.ac.be/synthesis/mbrola/
Elan Babeltech Infovox	Several well-established TTS systems now under a new commercial roof	Acapela Group www.acapela-group.com
RealSpeak Speechify Eloquence	Several well-established TTS systems now under a new commercial roof	ScanSoft Inc. www.scansoft.com
SVOX	TTS for server-based and embedded solutions	ETH Zürich/SVOX AG Zürich www.svox.ch
DRESS	Different unit systems, see Table 14.5. Also available as embedded version, footprint below 1 MB	Technische Universität Dresden www.ias.et.tu-dresden.de/sprache
Logox	TTS system with small databases using so-called microsegments (380 microsegments for one language)	GData Software AG www.logox.de

- The output of measuring devices, such as thermometers and kitchen scales.
- Employment related equipment, such as that used by physiotherapists. This is particularly useful, as physiotherapy is a typical profession for blind persons, but also has benefits for their sighted colleagues.
- Talking watches and calculators. The talking pocket calculator was a pioneering application of speech synthesis when pocket calculators were introduced into the classroom.

A range of household and other everyday living devices are discussed in more detail in Chapter 17 of this book.

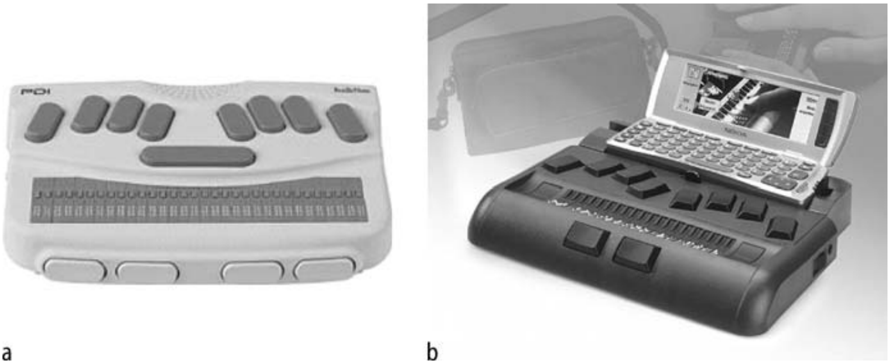


Figure 14.21a,b. Portable text-based devices: **a** BrailleNote BT 32 (by courtesy of Pulse Data); **b** the Nokia Communicator combined with Braillino (by courtesy of HandyTech)

14.6.3 Portable Text-based Devices

Due to the fast progress of computer technology, portable computing devices, which are variously called hand-helds, palmtops, organizers, pocket PCs or PDAs (personal digital assistants), have developed rapidly. Similar devices for blind users should combine the computing power and functionality of these devices with a Braille keyboard (with 8 keys to enable computer Braille) as input and a Braille display and/or a TTS system as output. Aspects of the design are discussed in detail by Bruynooghe (2001). There are two general approaches:

- The first type is a stand-alone device, as illustrated in Figure 14.21a. It has dedicated input and output and the same features as a modern pocket PC (Windows CE, e-mail, Word *etc.*). The data can be transferred to a desktop PC by means of a wireless interface.
- The second approach utilizes powerful standard devices which are combined with a special input/output device *via* a docking station. Figure 14.21b shows an example. This is an effective solution, as the basic device (the Nokia Communicator in this case) needs no modification. The available flash is utilized for the dedicated software without the expense of additional hardware.

The Braille display is generally optional. Features of the speech output should be controllable. This concerns mainly the speaking rate and (less important) other characteristics like pitch.

14.6.4 Access to Computers

The user interface must be complemented by appropriate software and hardware to allow blind and visually impaired users to access desktop or laptop computers. This is an application where speech input and/or output can be very effective and useful.

Output interface

In standard applications, the output of the user interface of a PC or workstation is mainly directed to the computer screen. Magnification of the screen contents, using software or sometimes hardware *magnifiers* can be used to give access to visually impaired users with sufficient vision. Magnifiers are described in Chapter 13. For other users, visual output must be replaced by audio and/or tactile output. The available software systems are generally able to control both a TTS system and a Braille display. There are two large groups of tasks which must be performed by the software systems:

- *Interaction with the operating system and application software*
  As indicated at the end of Section 14.5.2, the system must provide good quality support for screen navigation by the user to enable them to quickly obtain an overview of the screen contents, as well as for user interaction with the software. In addition, the contents of longer texts, for instance, in a text processor, must be accessible. A software system which supports these actions is called *screen reader*. The properties of screen readers are discussed in more detail in Chapter 13.
- *Accessing the Web*
  Speech and/or Braille output is generally produced from text. Documents with graphical layouts and/or a large number of pictures, which are found mainly on the Web, present two difficulties. These are difficulties in detecting the logical structure of the text, which is a prerequisite for reading it, and in transforming the graphical contents to speech or Braille, in order to make the contents comprehensible. Systems called *web analyzers* are used in conjunction with a web browser to present the contents of web sites in audio form. Website design for visually impaired users is discussed in Chapter 12.

It should be noted that there is a tendency towards the production of single unified systems including screen readers, web analyzers and magnifiers. The growing importance of portable computers means that screenreaders are also required for their operating systems, such as Windows CE or Windows for Pocket PC. For instance, a screenreader is required if a pocket PC is used together with a GPS receiver for navigation purposes (for details see Chapter 8).

A screen reader can be equipped with a standard application interface to allow the use of several TTS systems. This flexible solution allows the user to select the voice which they prefer, as well as to choose the language. The resulting additional possibilities for controlling the speaking rate, pitch and other parameters, give the user considerably scope in choosing the preferred voice.

Input interface

The keyboard is the standard computer input device. Blind users are able to use a standard keyboard if they can learn the arrangement of the keys for touch-typing, which is referred to as "typing blind" in languages such as German and French. This category of users will not require any additional input devices. However, since the visual information on the screen is not available to blind users, feedback to monitor their input and choice of operations is required from a Braille display or speech output.

Additional input devices are required by users who are unable to use a standard keyboard. Such devices include speech recognition systems, though the level of practical use is currently limited, but this could change in the future. Modern operating systems provide Voice-APIs (Application Programming Interfaces) which can be connected to speech recognition systems. In some cases, activation of integrated speech recognition software is the first step in speech input.

However, complete voice control of a computer involves considerably more than a limited number of input commands. This is particularly true for users who want to write their own programs, where the voice software must be able to engage in dialogue (Arnold *et al.* 2000). Further developments in this area are still required.

Similar problems are observed with dictation systems. A dictation system could be an extremely useful tool for a blind user, but word recognition is still not perfect. Therefore, correction of the converted text is required. Although current dictation systems provide powerful correction tools (including menus with proposals for substitutions), the use of these tools requires visual feedback, which is a significant barrier to their widespread use by blind and visually impaired people.

14.6.5 Reading Machines

Reading books and documents was mentioned in Section 14.5.2 when text-to-Braille conversion was introduced and Chapter 15 will provide a more general discussion. As can be seen in Figure 14.20, a TTS system can be included to provide a speech component.

This allows a reading system to be constructed from the following standard components: a flat-bed scanner, a PC, powerful OCR software, and a TTS system. In commercial systems the components are integrated and there are additional features, such as storage of the text files of scanned documents for later re-use. Despite the possibilities of simple solutions based on standard components, there are a number of commercially available dedicated reading machines, which are easier to use. An example is shown in Figure 14.22. In addition, insurance companies are more likely to recognise these specialised reading machines as a necessary aid than standard PC equipment.

Figure 14.22. Reading system Audiocharta compact (by courtesy of SilverCreations Software AG)

14.6.6 Access to Telecommunication Devices

There has been considerable development of telecommunication systems over the last decade and this has had an impact on telecommunication use by visually impaired and blind people. "Historically, people with a hearing disability have been the group facing the most problems when using telephones; however, with the ever increasing reliance on visual means for displaying information, it is increasingly visually impaired people who have been confronted with access problems" (Roe, 2001, p 30).

Speech technology can provide potential solutions, as in the case of the following input/output functions for mobile phones:

- Speech recognition is frequently used for voice-dialling. This feature was originally developed mainly for hands-free telephony in cars.
- Speech synthesis will be increasingly used for improving the user interface (speech MMI), caller name announcement, reading short messages (SMS) and remote access to e-mails.

Although these features were developed originally to provide improved performance for sighted users, they are very useful for visually impaired people and illustrate the benefits of a *design for all* approach. The technical prerequisites are the development of embedded speech input/output solutions (Hoffmann *et al.* 2004).

Despite the benefits of design for all, it is not able to resolve all problems and therefore visually impaired telecommunications users also require some special equipment. For instance, the Braillino system shown in Figure 14.21b is illustrated in combination with the Nokia Communicator. However, it can be used more generally with any mobile phone that uses the Symbian operating system, which is the global industry standard operating system for smartphones (www.symbian.com). This includes the Series 60 Phones (without an alphanumeric keyboard) and the Series 80 Phones (with an organizer function and an alphanumeric keyboard). The connection can be wireless *via* Bluetooth. From the functional point of view, the communication software (called Talks&Braille) acts as a screen reader for the Symbian operating system.

14.7 Discussion and the Future Outlook

14.7.1 End-user Studies

Potential users of speech technology would like to have (comparative) information on the performance of the available systems. However, it is difficult to obtain global comparative evaluations, due to the complexity of the systems and the fact that the evaluation criteria depend on the intended application. The studies carried out to date can be grouped and discussed as follows.

Evaluation of research systems

Progress in speech technology is normally measured in terms of improved word recognition rates (for recognizers) or improved scores when rating the naturalness (for synthesizers). Therefore, there are presentations giving an ongoing evaluation of research systems at the leading conferences. The availability of common databases allows the results of the evaluation of different systems to be compared. However, these research-oriented results relate to systems that are not yet commercially available, rather than the current state of the market.

Comparison with human performance

It is natural to compare speech technology with human performance. Every user of speech technology soon notices that it does not perform nearly as well as a person, but there are few quantitative assessments of this difference in performance. A fundamental investigation was carried out by Lippmann (1997) for speech recognizers. He demonstrated how the recognition rate breaks down in the presence of environmental noise, whereas human listeners perform essentially better. Corresponding results can be obtained by rating the quality of speech synthesis using a mean opinion score (MOS) scale ranging from 1 (bad) to 5 (excellent). The naturalness of human speech is rated close to 5, but the output of TTS systems is generally valued somewhere in the middle range; between 1.73 and 3.74 according to the survey by Alvarez and Huckvale (2002). Considerable further research will be required to close the gap in both speech recognition and speech synthesis compared to a human listener or speaker, respectively.

Evaluation of commercial systems

Before including speech technology in a product, a company generally evaluates a number of competing systems, though the results are only published occasionally. This type of study gives an interesting insight into the real performance of the available products. For example, Maase *et al.* (2003) investigated the performance of command and control speech recognizers for controlling a kitchen device. Usability studies showed that the users are accepting this kind of control for recognition rates greater than 85%. Tests with eight different products showed that this performance was never reached in real environments. Typical parts of these results are shown in Figure 14.23.

General product studies are very time-consuming and expensive. Therefore, they require a sponsor who has not got a vested interest in one of the products. For instance the study of ten different dictation systems (Flach *et al.* 2000) mentioned in Section 14.3.3 was originally produced for a computer journal. A more recent study (Stiftung Warentest 2004) of six dictation systems was carried out without publishing the recognition rates. The system with the best performance is indicated in the previously shown Table 14.8.

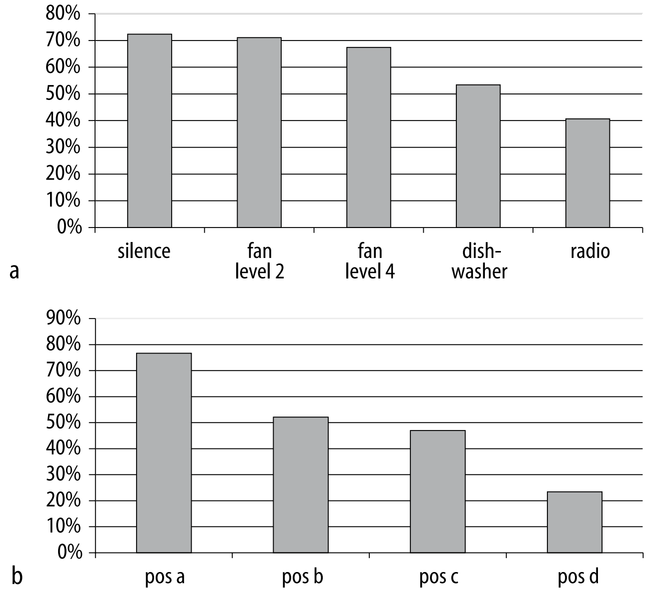


Figure 14.23a,b. Selected results from the study of Maase *et al.* (2003). The diagrams show the recognition rate of selected C&C recognizers for different noises (**a**) and different speaker positions (**b**). The speaker positions describe different places in the usability lab with growing distance (from 1 to 7 m). Reprinted by courtesy of the authors

Evaluation for user groups with special needs

There is clearly a need for studies of speech support and dictation systems for blind and visually impaired people. Unfortunately, there is a distinct lack of large scale user studies of speech support systems for this user group. However, there are several more general studies which include consideration of speech technology to a certain extent. A number of such investigations have considered improving learning environments for blind students (Kahlisch 1998). Another emerging field is the study of the assistive technology needs for elderly people. Since many elderly people have acquired visual impairments, these studies include useful material on speech-related technologies. Figure 14.24 presents an example.

14.7.2 Discussion and Issues Arising

An overview of the remarks in this chapter shows that the performance of speech input/output systems is by no means perfect, despite improved algorithms, larger databases, increased memories, and growing computing power. In general, this still somewhat disappointing performance is due to the extreme complexity of human speech processing which it is difficult to satisfactory approximate by technical systems. Although there is not space to discuss the reasons for this less than satisfactory performance in detail, some of the reasons for this are briefly summa-

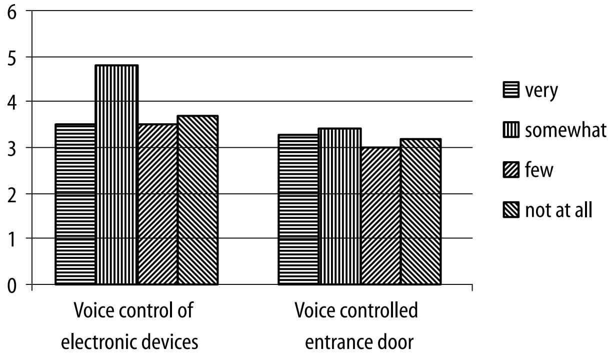


Figure 14.24. Example of a usability study. The diagram shows the acceptance of speech controlled services by different user groups according to the study of Hampicke (2004). Reprinted by courtesy of the author. Score of 6: in any case. Score of 3: medium. Score of 0: in no case. The legend describes the grade of visual impairment

rized in Table 14.10. Examining this table leads to the following conclusions about important future directions for basic research in speech technology:

- Speech understanding.
- Acoustic front end.
- Modelling human speech and language processing.

These topics are all highly interdisciplinary and will require interdisciplinary work.

14.7.3 Future Developments

As discussed in this chapter, speech technology has established itself as a stable and successful component of assistive technology. Speech technology is also becoming increasingly successful in other fields with a greater economic impact, including in the telecommunications area for communication with call centres and telephone banking. Although beyond the remit of this chapter, a survey of user opinions of this technology would be interesting, since there is at least anecdotal evidence that users prefer to communicate with a person and are highly dissatisfied with call centres. According to recent data (Sohn 2004), the turnover in business applications of speech technology will grow from $ 540 millions currently worldwide to $ 1600 millions in the year 2007.

This growth in the use of speech technology is not surprising in view of the importance of speech in telephone applications and consequently also for their automation. The importance of speech input/output systems relative to other media is likely to grow, as can be seen from Table 14.11.

What will this tendency mean for blind and visually impaired people? Developments in speech technology will improve access to interfaces for an increasing range of applications for this group (though not for deafblind people). The resulting benefits are likely to be substantial and cover applications ranging from access to numerous knowledge sources to improved accessibility of household appliances.

Table 14.10. Actual research problems in speech technology, explained by means of the general scheme of a speech processing system (Figure 14.4)

Where are the problems localized in Figure 14.4?	How can the problems be described?	What can research do to solve the problems?	Examples for first solutions
At the top of the figure	Our systems do not *understand* what they do. The scheme is ending at text level without semantic components	Develop speech understanding, cooperate with computer linguistics/AI/semiotics	Speech-to-speech translation systems like Verbmobil (Wahlster 2000) In speech synthesis: concept-to-speech (CTS) instead of TTS
At the bottom of the figure	The acoustic channel between the user and the converters (microphone or loudspeaker, respectively) is still neglected in most cases	Consider the system (recognizer or synthesizer, respectively) and the acoustic environment as a unit and develop the "acoustic frontend"	Acoustic signal processing such as: • Microphone arrays. • Noise suppression. • Source separation. • Directed sound supply.
In the components of the figure	Because our understanding of human speech processing is far from an applicable level, the models which we use are more or less mathematical or empiric	Although a technical system needs not to be a close copy of the biological counterpart, we need essentially more knowledge of human speech production and perception	Many activities in modelling prosody in close cooperation of engineers and phoneticians during the last decade; Research systems which model human acoustic processing

Table 14.11. How to interact with future systems? An overview from Weyrich (2003)

Small devices	Speech
Service robots	Speech and gestures, artifical skin, emotions
Federation of systems	Speech and gestures, emotions
e-Business	Active dialogue systems, interactive multimedia
Augmented reality systems	Speech, gestures

Talking products which are of interest to both sighted and visually impaired people are more attractive to companies due to their larger markets and therefore this type of product is more likely to be widely available from standard suppliers and at a reasonable price than specialised products for visually impaired people. For instance, blind and many visually impaired users require speech (or tactile) output to state the function of the key being pressed or the knob setting on the (complex) control panel of a washing machine. This audio option may also be of interest to sighted users. The inclusion of both speech and tactile output could be considered part of a design for all approach, but, as already indicated, though design for all should be part of good design practice, it will never totally replace the need for assistive devices.

There is therefore considerable potential for increasing accessibility to blind and visually impaired people, though further technical developments will be required. However, it should also be noted that access to new technologies is limited by a number of different factors, including geography and poverty. The term 'digital divide' is often used to describe the difference between people who do and do not have access to modern technologies and the resulting disadvantages, whereas the term eInclusion is used for access to the information society by disabled people and other potentially disadvantaged groups. While it is important to ensure that blind and visually impaired people are able to fully participate in the information society, it should also be recognised that some people, both blind and sighted, do not like technology. It will therefore also be important to ensure that there are also low technology accessibility solutions for blind and visually impaired people and that information is available in a number of different formats, including but not solely electronically.

Speech and language technology will always be compared to natural human speech and language. Therefore, regardless of progress, they are likely to be found wanting for a long time to come, if not permanently. This presents an ongoing challenge, which is probably much greater than that encountered in many other disciplines. As Waibel and Lee (1990) state in their preface to *Readings in Speech Recognition*: "Many advances have been made during these past decades; but every new technique and every solved puzzle opens a host of new questions and points us in new directions. Indeed, speech is such an intimate expression of our humanity—of our thoughts and emotions—that speech recognition is likely to remain an intellectual frontier as long as we search for a deeper understanding of ourselves in general, and intelligent behaviour in particular."

Acknowledgement. As can be seen from the list of references, the material in this chapter is based on research results and teaching material of the chair for speech communication at the Technische Universität Dresden. The author would like to take the opportunity to thank his team for their fruitful cooperation on many projects.

Special thanks for helpful discussions and support to Professor Dieter Mehnert, formerly at the Humboldt-Universität zu Berlin, Professor Klaus Fellbaum, Brandenburgische Technische Universität Cottbus, Professor Gerhard Weber, Universität Kiel, and Dr. Lothar Seveke, Computer für Behinderte GmbH, Dresden.

Questions

Q.1 Give the following acronyms in full: TTB, STT, BTT, TTS.
Q.2 Explain the differences between narrowband and broadband spectrograms.
Q.3 Explain the differences between numerical and structural pattern recognition.
Q.4 Explain the steps which are required to convert speech into text.
Q.5 Explain the steps which are required to convert text into speech.
Q.6 Compare the application of speech and Braille output in assistive devices with respect to their benefits and drawbacks.

Projects

P.1 Project in speech analysis.

Practical work in speech technology requires basic knowledge of and active experience with speech signals. Therefore, the aim of this first project is an investigation of the properties of real speech and, in particular, the study of different types of sounds. The following classes of speech sounds should be considered:

- Vowels and diphthongs.
- Voiceless stops (/p/, /t/, /k/).
- Voiced stops (/b/, /d/, /g/).
- Voiceless fricatives (/f/, /θ/, /s/, /ʃ/, /h/).
- Voiced fricatives (/v/, /ð/, /z/, /ʒ/).
- Nasals (/m/, /n/, /ŋ/).
- Liquids (/r/, /l/).
- Glides (/ω/, /j/).

Signals should be recorded using a microphone and PC soundcard. Both single sounds and complete words or phrases should be recorded to allow you to investigate the differences between isolated sounds and fluent speech in both the time and frequency domains. You will require software that is able to perform the steps presented in Section 14.2.1 and to view and compare speech waveforms and spectrograms. Additional information can be found in textbooks on experimental phonetics, such as Olive *et al.* (1993).

Signal analysis software which can be found on the Web includes the following:

- Speech filing system (SFS) has many tools for speech related signal processing. It was developed by Mark Huckvale and can be obtained *via* the web site http://www.phon.ucl.ac.uk/.
- WaveSurfer is an open source tool for sound visualization and manipulation and can be downloaded using the web site http://www.speech.kth.se/wavesurfer/.

- Praat features many tools and covers advanced algorithms for labelling and speech synthesis. It can be obtained *via* http://www.praat.org/ or http://www.fon.hum.uva.nl/praat/.

Some textbooks on signal processing, such as Karrenberg (2002) (English version) and Karrenberg (2004) (German version) have companion CD-ROMs which include a suite of signal analysis software. This approach has the advantage of providing additional information on digital signal processing.

P.2 Project in sound recognition.

Speech recognizers (see Section 14.3) are complex software systems which require considerable experience to be designed appropriately. To function satisfactorily powerful recognition systems require a considerable volume of training data which is often difficult to collect and label. This project will produce a simplified recognizer which is still able to demonstrate the main components of a recognition system. The system should be able to distinguish between two isolated long (static) vowels, such as: /a:/ (like f***a***ther in English or St***aa***t in German) and /i:/ (like b***ea***d in English or K***ie***l in German).

As shown in Figure 14.7, static vowels can be characterized by means of the formant frequencies F1 and F2. This property was used in the examples in Section 14.3 (see Figure 14.8). The formant frequencies could be estimated from the spectrum. However, in simple cases such as the separation of two vowels, the spectral analysis can be replaced by a zero-crossing analysis of the waveform which can be implemented more easily.

The zero-crossing rate is the number of zero-crossings of the signal in a time window of given length (see Figure 14.25). It is easy to detect a zero-crossing in digital signal processing because it corresponds to the change of the sign of two neighbouring samples. To reduce the influence of noise, the signal should be allowed to reach a given threshold value after crossing the time axis, before subsequent zero-crossings are counted.

The waveform of the differentiated signal of the static vowels is also required. In digital signal processing, a digital signal $x(k)$ is differentiated by forming a new series of samples $\Delta x(k)$, which are calculated as the difference:

$$\Delta x(k) = x(k) - x(k-1) \tag{14.10}$$

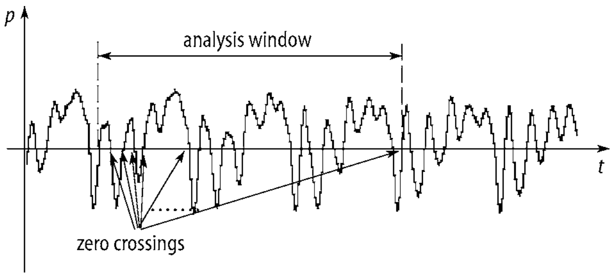


Figure 14.25. Zero-crossing analysis of a speech signal (from Kordon and Kürbis 2004)

A signal distance classifier can now be constructed as follows:

1. Record a vowel /a:/ or /i:/ and select a time window of suitable length. It is recommended to select a window which has sufficient signal energy.
2. Calculate the zero-crossing rate of the recorded signal.
3. Calculate the derivative of the recorded signal using Equation 14.10.
4. Calculate the zero-crossing rate of the differentiated signal.
5. Form the feature vector describing the vowel /a:/ or /i:/ using the zero-crossing rates obtained in steps 2 and 4.
6. Repeat steps 1–5 until you have a sufficient data for both /a:/ and /i:/.
7. From this data, compute the means of the vectors for the two classes separately. These means will be used as the representatives $\boldsymbol{r}_\mathrm{a}$ and $\boldsymbol{r}_\mathrm{i}$ of the two classes, as discussed in Section 14.3.1.
8. Write an algorithm which calculates the distances $d(\boldsymbol{r}_\mathrm{a}, \boldsymbol{x})$ and $d(\boldsymbol{r}_\mathrm{i}, \boldsymbol{x})$ between a given two-dimensional feature vector $\boldsymbol{x}$ and the representatives computed in step 7.

All these steps could be performed manually. However, there are advantages in writing a computer program to implement them. It should now be possible to classify an "unknown" sound as /a:/ or /i:/ using the decision rule given in Equation 14.5.

P.3 Project in concatenative speech synthesis.

Experiments in diphone concatenation can be used to improve understanding of the problems in the acoustic component of a TTS system. However, the preparation of a complete diphone set would require excessive effort. Therefore the tutorial "Speech Synthesis" (Hoffmann *et al.* 1999b) is recommended instead. It can be found at the following URL:

http://www.ias.et.tu-dresden.de/sprache/lehre/multimedia/tutorial/index.html

One section of this tutorial offers a selection of carrier words. You should define the boundaries of the diphones contained in these words in an interactive way and combine these diphones in an arbitrary way to form new words.

References

Alvarez, Y., and Huckvale, M., 2002, The reliability of the ITU-T P.85 standard for the evaluation of text-to-speech systems. In: *Proc. Int. Conf. on Spoken Language Processing (ICSLP),* Denver, Vol. 1, pp. 329–332

Antonacopoulos, A., 1997, Automatic reading of Braille documents, In: Bunke, H., Wang, P.S.P. (Eds.). *Handbook of Optical Character Recognition and Document Image Analysis.* World Scientific Publishing Co., Singapore, pp. 703–728

Arnold, S.C., Mark, L., and Goldthwaite, J., 2000, Programming by voice, VocalProgramming. In: *Proc. The Fourth International ACM SIGCAPH Conference on Assistive Technologies (ASSETS'00)*, Arlington, VA, pp. 149–155

Bruynooghe, J., 2001, EuroScope: Development of a wireless portable Braille organiser for visually impaired. In: *Proc. Conf. and Workshop on Assistive Technologies for Vision and Hearing Impairment (CVHI 2001)*, Castelvecchio Pascoli, Italy, 7 p.

Bunke, H., and Wang, P.S.P. (Eds.), 1997. *Handbook of Optical Character Recognition and Document Image Analysis*. World Scientific Publishing Co., Singapore

Chu, W.C., 2003, *Speech coding algorithms. Foundation and evolution of standardized coders.* Wiley-Interscience, Hoboken, New Jersey

Deller, J.R., Proakis, J.G., and Hansen, J.H.L., 1993. *Discrete-Time Processing of Speech Signals*. Macmillan, New York. Reprint: John Wiley & Sons 1999

Dutoit, T., 1997. *An Introduction to Text-to-Speech Synthesis*. Kluwer, Dordrecht

Eichner, M., Wolff, M., and Hoffmann, R., 2000, A unified approach for speech synthesis and speech recognition using stochastic Markov graphs. In: *Proc. 6th Int. Conf. Spoken Language Processing (ICSLP)*, Beijing, China, Vol. 1, pp. 701–704

Eichner, M., Wolff, M., and Hoffmann, R., 2004, Voice characteristics conversion for TTS using reverse VTLN, In: *Proc. Int. Conf. in Acoustics, Speech, and Signal Processing (ICASSP),* Montreal, Canada, Vol. 1, pp. 17–20

Fant, G., 1960, *Acoustic Theory of Speech Production,* Mouton & Co., 's-Gravenhage

Fellbaum, K., 2000, Elektronische Sprachsignalverarbeitung – Stand der Technik, Zukunftsperspektiven. In: *Proc. Elfte Konferenz Elektronische Sprachsignalverarbeitung*, Cottbus, Germany. Studientexte zur Sprachkommunikation, Vol. 20, pp. 9–16

Flach, G., Hoffmann, R., and Rudolph, T., 2000, Eine aktuelle Evaluation kommerzieller Diktiersysteme. In: *Proc. 5. Konferenz zur Verarbeitung natürlicher Sprache (KONVENS)*, Ilmenau, Germany. ITG-Fachbericht 161, VDE-Verlag, Berlin, pp. 51–55

Fujisaki, H., 1997, Prosody, models, and spontaneous speech, In: Y. Sagisaka *et al.* (Eds.), *Computing Prosody*. Springer, New York, pp. 27–42.

Görz, G., Rollinger, C.-R., and Schneeberger, J., 2000, *Handbuch der Künstlichen Intelligenz.* Third edition, Oldenbourg Verlag, München und Wien

Gray, L., 1998, *Braille Transcribing*. http://www.uronramp.net/~lizgray/ (Data accessed June 6, 2004)

Hersh, M., and Johnson, M. (Eds.), 2003. *Assistive Technology for the Hearing-impaired, Deaf and Deafblind*. Springer, London

Hess, W., 2003, Articulatory, parametric, concatenative – Three paradigms in acoustic synthesis of speech and their implications. In: Wolf, D. (Ed.), *Signaltheorie und Signalverarbeitung (Festschrift for A. Lacroix)*. Studientexte zur Sprachkommunikation, Vol. 29, pp. 116–125

Hoffmann, R., 1998, *Signalanalyse und -erkennung*. Springer-Verlag, Berlin

Hoffmann, R., Hirschfeld, D., Jokisch, O., Kordon, U., Mixdorff, H., and Mehnert, D., 1999a, Evaluation of a multilingual TTS system with respect to the prosodic quality, In: *Proc. Intern. Congress of Phonetic Sciences (ICPhS)*, San Francisco, CA, Vol. 3, pp. 2307–2310

Hoffmann, R., Ketzmerick, B., Kordon, U., and Kürbis, S., 1999b, An interactive tutorial on text-to-speech synthesis from diphones in time domain. In: *Proc. 6th European Conf. on Speech Communication and Technology (EUROSPEECH)*, Budapest, Hungary, Vol. 2, pp. 639–642

Hoffmann, R., Jokisch, O., Hirschfeld, D., Strecha, G., Kruschke, H., Kordon, U., and Koloska, U., 2003, A multilingual TTS system with less than 1 Mbyte footprint for embedded applications. In: *Proc. Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP),* Hong Kong, Vol. 1, pp. 532–535

Hoffmann, R., Jokisch, O., Strecha, G., and Hirschfeld, D., 2004, Advances in speech technology for embedded systems. In: *Proc. Conf. and Workshop on Assistive Technologies for Vision and Hearing Impairment (CVHI 2004)*, Granada, Spain, 6 p.

Kahlisch, T., 1998, *Software-ergonomische Aspekte der Studierumgebung blinder Menschen*. Kovac, Hamburg

Karrenberg, U., 2002. *An Interactive Multimedia Introduction to Signal Processing*. Springer-Verlag, Berlin

Karrenberg, U., 2004. *Signale – Prozesse – Systeme. Eine multimediale und interaktive Einführung in die Signalverarbeitung.* Third edition, Springer-Verlag, Berlin

Kempelen, W. von, 1791, *Mechanismus der menschlichen Sprache nebst der Beschreibung seiner sprechenden Maschine*. J.W. Degen, Wien

Kordon, U., and Kürbis, S., 2004, *Akustischer Schalter*. Praktikumsanleitung, TU Dresden

Levelt, W.J.M., 1989. *Speaking: From Intention to Articulation*. The MIT Press, Cambridge, Mass

Levinson, S.E., 1985, Structural Methods in Automatic Speech Recognition, In: *Proceedings of the IEEE*, Vol. 73, No. 11, pp. 1625–1650

Lippmann, R.P., 1997, Speech recognition by machines and humans. In: *Speech Communication*, Vol. 22, No. 1, pp. 1–16

Maase, J., Hirschfeld, D., Koloska, U., Westfeld, T., and Helbig, J., 2003, Towards an evaluation standard for speech control concepts in real-world scenarios. In: *Proc. 8th European Conf. on Speech Communication and Technology (EUROSPEECH)*, Geneva, Switzerland, pp. 1553–1556

Olive, J.P., Greenwood, A., and Coleman, J., 1993. *Acoustics of American English Speech. A Dynamic Approach*. Springer-Verlag, New York

Rabiner, L., and Juang, B.-H., 1993. *Fundamentals of Speech Recognition*. Prentice Hall, Englewood Cliffs, NJ

Roe, P.R.W. (ed.), 2001, *Bridging the gap? Access to telecommunications for all people*. Commission of European Communities, COST 219bis, Lausanne

Schmidt, K.O., 1932, *Verfahren zur besseren Ausnutzung des Übertragungsweges*. D.R.P. Nr. 594976

Schnell, M., Jokisch, O., Küstner, M., and Hoffmann, R., 2002, Text-to-speech for low-resource systems. In: *Proc. on the 5th IEEE Workshop on Multimedia Signal Processing (MMSP)*, St. Thomas, Virgin Islands, 4 pp.

Schutkowski, G., 1938, *Optisch-elektrische Einrichtung für Blinde zum Umwandeln von Schriftzeichen in Blindenschrift*. D.R.P. Nr. 657536

Schutkowski, G., 1952. Druckschriftlesemaschine für Blinde. In: *Die Technik*, Berlin, Vol. 7, pp. 266–270 and pp. 335–337

Shannon, C., 1949. Communication in the presence of noise, Reprinted in: *Collected Works*, ed. by N.J.A. Sloane and A.D. Wyner, New York: IEEE Press 1993, pp. 160–172

Sohn, G., 2004, *Spracherkennung und Sprachautomatisierung boomen*. Pressetext Deutschland, June 26

Stiftung Warentest, 2004, Ein eheähnliches Verhältnis. In: *test*, Vol. 39, No. 1, pp. 38–41.

Wahlster, W. (Ed.), 2000. *Verbmobil: Foundations of Speech-to-Speech Translation*. Springer Verlag, Berlin

Waibel, A., Lee, K.-F. (Eds.), 1990. *Readings in Speech Recognition*. Morgan Kaufmann Publishers, San Mateo, CA

Weber, G., 2004, Empowering IT for social integration. In: *Proc. Conf. and Workshop on Assistive Technologies for Vision and Hearing Impairment (CVHI 2004)*, Granada, Spain

Werner, S., Eichner, M., Wolff, M., and Hoffmann, R., 2004, Towards spontaneous speech synthesis—Utilizing language model information in TTS. In: *IEEE Trans. on Speech and Audio Processing*, Vol. 12, No. 4, pp. 436–445

Weyrich, C., 2003, Industrial view on Human Computer Interaction, In: *Proc. Human Computer Interaction Status Conference*, Berlin, Germany. German Federal Ministry of Education and Research, pp. 13–18.

Further Reading

Books

Furui, S., 2001: *Digital Speech Processing, Synthesis, and Recognition*. Second edition, Marcel Dekker, New York

Holmes, J.N., 1988. *Speech synthesis and Recognition*. London: Chapman and Hall. German translation by G. Ruske published in 1991, Oldenbourg Verlag, München und Wien

Jurafsky, D., and Martin, J.H., 2000. *Speech and Language Processing*. Prentice Hall, New Jersey

Selected Periodicals in Speech Technology

IEEE Transactions on Speech and Audio Processing. A publication of the IEEE Signal Processing Society. IEEE, New York

Speech Communication. A publication of the European Association for Signal Processing (EURASIP) and of the International Speech Communication Association (ISCA). Elsevier Science B. V., Amsterdam

Computer Speech and Language. Academic Press/Elsevier, Amsterdam *etc*

International Journal of Speech Technology. Published in association with the Applied Voice Input/Output Society (AVIOS). Kluwer Academic Publishers, Norwell, MA

Leading Conferences in Speech Technology

EUROSPEECH (European Conference on Speech Communication and Technology). An Interspeech event
ICSLP (International Conference on Spoken Language Processing). An Interspeech Event
ICASSP (IEEE International Conference on Acoustics, Speech, and Signal Processing)
ICPhS (International Congress on Phonetic Sciences)

Resources

The following web sites give information on general aspects of speech and language technology. For further details, see the URLs cited in the text and the tables. Although the author considers the external web sites referred to in this chapter useful, he does not vouch for their accuracy.

collate.dfki.de	COLLATE – Computational Linguistics and Language Technology for Real Life Applications (includes the Virtual Information Center Language Technology World)
www.elsnet.org/	ELSNET – European Network of Excellence in Human Language Technologies
www.hltcentral.org/htmlengine.shtml?id=165	JEWELS – Joint European Website for Education in Language and Speech
www.isca-speech.org	ISCA – International Speech Communication Association
www.phon.ucl.ac.uk/	UCL London, Department of Phonetics and Linguistics
www.elra.info/	ELRA – European Language Resources Association
www.ecess.org	ECESS – European Center of Excellence on Speech Synthesis
ttssamples.syntheticspeech.de/	Overview on many TTS systems including demonstrations for the German language
mambo.ucsc.edu/	Perceptual Science Laboratory at the University of California
www.cslu.ogi.edu/	Center for Spoken Language Understanding at the Oregon Graduate Institute of Science and Technology
db1.rehadat.de/rehadat/alliance.jsp	International Alliance of Assistive Technology Information Providers

15 Accessing Books and Documents

Learning Objectives

Reading is an essential daily living task, and is crucial for school and work. Whether it is sorting the bills, reading a textbook or the daily newspaper, access to reading is critically important to people with disabilities that prevent easy reading of the printed page. Assistive technology has been created to address these needs and bridge the accessibility challenge to print. This chapter presents the wide array of these solutions, and explains the underlying technology. All of these solutions have the common challenge of acquiring the text to be made accessible, and the choice of how to present the accessible text. Text acquisition can be as varied as scanning the printed page and doing optical character recognition to directly downloading the text from the Internet. Accessible presentations range from having a human reader narrate the text, to enlarged print, to Braille to synthetic speech. This chapter on print accessibility has the following learning objectives:

- Understanding the general principles of reading systems.
- Optical character recognition technology and its application to reading systems.
- Learning about talking books, including the new Digital Accessible Information SYstem and the devices that play these audio books.
- Exploring access to textbooks and daily newspapers.
- Projecting potential future developments in the field.

15.1 Introduction: The Challenge of Accessing the Printed Page

"*All my life, I have wanted to read a book and drive a car*", said one of the first blind people I met. Access to the printed word has traditionally been the primary focus of adaptive technology for the blind and visually impaired, and has been technically more practical than mobility solutions. Print access is central to educational and vocational opportunity, as well as participation in civil and social institutions. The challenge is to change the inaccessible book, newspaper or document into a form that can be used by a person with a vision disability.

Magnification is the traditional method for addressing vision loss and access to text. First accomplished with optical lenses, and later with electronic closed circuit television (CCTV) devices, magnification uses a visually impaired person's remaining vision to provide access to text. Magnification cannot help individuals who are completely blind.

Braille was probably the most significant adaptive technology advance for the blind of the 1800s. The transformation of the printed word into tactile dots was a major breakthrough. It also heralded the concept of alternative techniques for the blind, where the lack of vision was not a barrier to accomplishing an important task such as reading.

When Thomas Edison invented the phonograph later in the nineteenth century, one of his first ideas for an application of his new invention was to create talking books for the blind. It was not until the 1930s that such a system was actually created at the US Library of Congress.

These three alternative techniques for accessing print, magnification, tactile Braille and audible speech, are at the core of almost all book and document access technology for the visually impaired [the Optacon used tactile recognition, but is no longer manufactured]. They are not necessarily exclusive: individuals frequently use two or even all three of these approaches.

This chapter will not focus on optical or CCTV magnification of books and documents; however, it will discuss additional visual technologies that extend what is possible with remaining vision. Our task will be to examine how documents are made accessible when the document itself needs to be transformed.

These transformations in the past have required human intervention. Generally, a sighted person reworks the document into an accessible form. The original approach was to have the sighted person read aloud to the visually impaired person, a technique still in wide use today. The Perkins Brailler and Braille printing presses are important tools for professionals to use to create Braille books. Human narrated books are widely available on audio cassettes.

Technology in use today has greatly expanded the options available for accessible reading and lessened the need to have a sighted person intervene in the process. We now have Braille transcription software, personal Braille embossers and refreshable electronic Braille displays. For audio, we have computer synthesized voices to speak aloud digital text (also known as Text-to-Speech, or TTS). With reading systems that use optical character recognition (OCR), we can provide access to Braille, audio and customized visual displays directly from the printed page. Reading systems now put tools in the hands of the person with a disability, making it possible for him or her to read independently without requiring the intervention of a sighted person.

The most common methods of accessing printed documents are described in Figure 15.1. Printed documents are transformed by different actors and methods into forms that the person with a visual impairment can read. We will be concentrating in this chapter on the methods that provide access through the use of technological tools.

These tools have revolutionized the access to printed material for people with vision impairments. Far more material is now available in accessible form, at ever

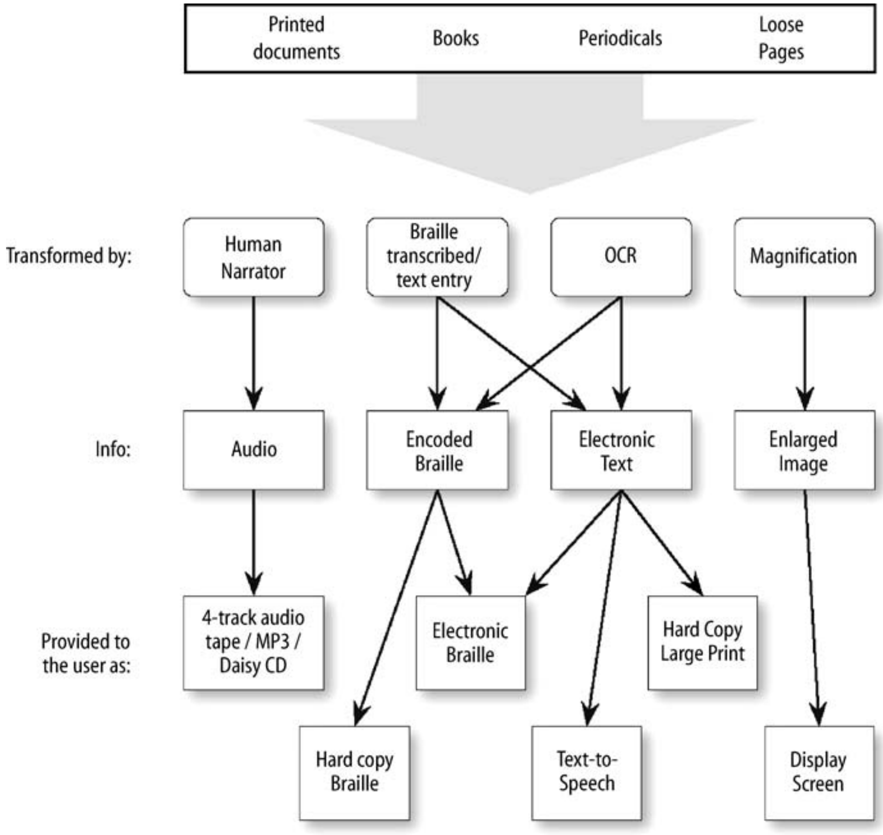


Figure 15.1. Printed document access challenge

decreasing costs, compared to past decades. The barriers to print access are coming down, and the technology roadmap for future advances is reasonably clear. More and more visually impaired people are driving their own access to information. Let us explore more about the technology that brings reading and independent literacy to people with visual disabilities.

15.2 Basics of Optical Character Recognition Technology

Delivering the actual text and structure of a document is increasingly the core requirement for text access. The process of having a sighted person intervene, by reading aloud books, newspapers and documents, or entering them as Braille, is on the wane because of expense and the delay in production. It is ironic that, in a world where almost all documents are generated electronically, access to those electronic documents is so difficult! Authors and publishers of books are concerned about piracy and worry about making books easily available in electronic form, although they rarely object to access for people with disabilities. Making the connection to the right person is logistically difficult. Solutions are needed to provide access directly to books.

Optical character recognition (OCR) systems meet this need. They recreate the document in electronic form from a scanned image of the document. A desktop image scanner is used to optically scan each page: taking a digital picture of the page.

The scanner is increasingly a standard piece of equipment in offices and home computing setups. A very popular product is the combined scanner/printer, which can also act as a copier. The most frequent configuration is a flat glass platen on which the document or book is placed. A scanning bar with a lamp moves under the glass plate, illuminating and imaging the page line by line, typically representing the image as an array of pixels at a resolution of 300 dots per inch (118 dots per centimetre). This scanned image file is transferred to a computer for additional processing. The scanned image of the page is quite useful for tasks such as sending a facsimile or making a photocopy.

However, the image cannot be directly used to generate Braille or synthetic voice output. OCR systems analyze the picture of the page, find the letters and words, and generate a text file with words, paragraphs and pages. This text file can be turned into Braille or sent to a voice synthesizer, and thus made accessible.

Imagine a business letter. A photograph of that letter can be used to reproduce the letter on a printed page or on a screen. However, the picture of the page will not be accessible to a blind person at all. Seeing the picture of the words is not the same as understanding the words, and a computer is able to communicate the words to a blind person only after the computer has identified the words within the picture.

OCR technology turns that picture into words (see Figure 15.2). On a personal computer a word processing program can then edit the letter, just as if someone had retyped that business letter. Because the words are available in the word processing program, a specialized computer program called a screen reader can read them aloud or send them to a Braille display or printer.

The OCR process often makes mistakes. Identifying the words from a picture of a page can be difficult, especially if the document is of poor quality or the print is small. The OCR technology breaks the picture of the page down into lines of text, and then further subdivides the picture into words and letters. By analyzing the picture of a letter and where it stands on the line, the OCR can usually tell which letter it is (for example, a capital 'P' *vs* a lower case 'p'). The OCR then builds up words, lines, paragraphs and pages. Understanding how the OCR process operates can help in recognizing its limitations.

15.2.1 Details of Optical Character Recognition Technology

OCR technology imitates the human perception process of visual reading. This technology has steadily progressed over the last 50 years, but it is still not the equal of human readers. OCR uses a methodical approach to analyzing what is on a page. A typical OCR process involves the following steps:

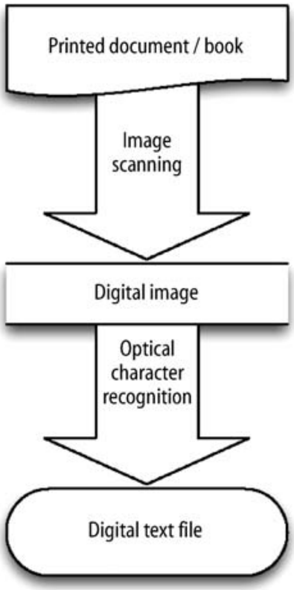


Figure 15.2. OCR process

Step 1. Adjusting the contrast of the page image
Step 2. Removing speckle and other image noise from the image of the page
Step 3. Identifying if the text is sideways or upside down (or if the page is blank)
Step 4. De-skewing the page (straightening it if it was scanned at a slight angle)
Step 5. Finding blocks of text on the page
Step 6. Finding the lines in the text blocks
Step 7. Identifying the baseline of the line of text (to distinguish between capital and lower-case letters)
Step 8. Isolating the pictures of individual words and letters
Step 9. Recognizing the letters by analyzing their features
Step 10. Assembling words from the recognized letters
Step 11. Resolving uncertain letters or words, using linguistic rules and dictionaries
Step 12. Reconstructing the lines, paragraphs and page in the desired format (for example, ASCII, Microsoft Word or RTF)

It is interesting to note that Step 3 above, where the orientation of the page is detected by the computer, was not a standard component of early commercial OCR systems because it was assumed that a person placing the page on the scanner could see whether or not the text was upside down before scanning. After being made aware of this need and seeing early implementations of it by organizations making reading systems for the blind, the OCR vendors added this accommodation so that an image scanned using the wrong orientation would be detected and digitally rotated into the correct orientation before the OCR proceeded. Now, this is a standard feature of all the commercial OCR packages, and sighted people benefit from it as well since they can now place a page on a scanner without checking that it is in the "correct" orientation.

Using the old computing dictum of "*garbage in, garbage out*" (also known as "*rubbish in, rubbish out*"), poor quality images will lead to inaccurate OCR. One of the major technology efforts in character recognition is improving the quality of the image being recognized. Good OCR depends on a good quality scanner to capture an accurate image of the text. Cameras, hand-operated scanners and inexpensive sheet-fed scanners (where the page is moved past a fixed line sensor) generally fall short of providing the best quality images for OCR. Flatbed scanners and higher end sheet-fed scanners reliably provide good page scans.

The standard image scanning resolution of 300 dpi/118 dpc is almost always sufficient for standard text documents. A typical small character scanned at this resolution is roughly 20–30 pixels high, which is enough for the OCR to distinguish similar looking characters (such as '*e*' *vs* '*c*'). Figure 15.3 shows examples of individual character images of varying quality.

Processing of the scanned image before recognition is critical to remove the garbage. Contrast is a critical parameter, especially in modern documents (such as magazines) with text printed on complex, and often multicoloured, backgrounds. Character recognition engines need just the characters and reject the background and picture content. This need competes with other imaging tasks where the goal is to accurately render the original document on the computer display. Some scanning processes have a binary contrast set manually, where others add automatic contrast technology or will process a colour or grey scale image of the document. Binary document images keep just one bit of data per pixel: whether it is black or white. Grey images typically keep 8 bits of data per pixel, which is equivalent to 255 values from black to white. Colour images are often scanned at 24 bits per pixel, 8 bits for each of 3 colours. For OCR, these are generally reduced to binary images through processing.

The other image processing steps in OCR are designed to make the image better after scanning. Despeckling removes noise: the little flecks of ink, paper imperfections or scanning quirks that otherwise might show up as punctuation in odd places on the reproduced pages. Orienting and straightening the page improves the accuracy of the OCR, and recognition of speciality font characteristics such as italics. Some OCR software can recognize the difference between blocks of text and blocks of image on a page, and do the appropriate processing on each type of block.

The middle part of the process is actually the recognition of a character or glyph. For example, in typography the letters 'fl' are typically printed as a single digraph rather than as two separate letters, so the OCR treats this as another type of character to be recognized. By processing the text blocks, lines of text and words in successively smaller units, the OCR engine is presented with a single character image to recognize. It generally does this by transforming the character image into a set of features. In the original OCR engines, this feature set might be as simple as normalizing the image onto a 5×9 grid and using which grid elements were on or off as the features. A very typical feature used in modern OCR is aspect ratio: the ratio of the height to the width of a character. For example, in many feature sets

Figure 15.3. Individual character images of varying quality

that do image normalization the only difference between a period and a vertical bar (which could be the letter 'l', the numeral '1' or the bar character itself '|') is the aspect ratio calculated before fitting the character onto a normalized grid.

Sometimes the differences between different characters can be quite small. The difference in a small point-size font between a small letter '*e*' and a '*c*' turns on a few pixels: if the contrast is set too lightly and the font is small, these differences might not be captured. As noted above, in some fonts there are no differences in appearance between the letter 'l' and the digit '1', or the letter 'o' and the digit '0', so these tend to be grouped as one recognized character.

Contextual analysis is used to separate these similar looking characters. For example, "look" can be easily told to be all letters, whereas "1004" is clearly all numbers. "Thc" is probably wrong and can be changed to "The" with confidence that this is almost always correct. If the word "return" is printed with the final 'r' and the 'n' touching, it may look like the word "retum". OCR systems use dictionary checks to see what words are permitted in a given language, turning "retum" into "return," but this is not always possible. For example, "modem" and "modern" are both allowable words in English. So, it is usually easy to tell when a book has been scanned using OCR, because the OCR makes mistakes that humans rarely would.

Since humans can read text with many errors and still understand what the correct text should be (an example is the deliberate misspelling in many current e-mail spam messages), we know that it is possible to solve this problem better. This is an area where continuing to enhance the contextual and document understanding technology would make character recognition much more accurate.

15.2.2 Practical Issues with Optical Character Recognition Technology

Users rate OCR technology on accuracy and speed. How good a job did the OCR do in recreating the page? Ideally, this is as similar as possible to the author's original electronic document. How long a wait is there from scanning to reading a given page? As short a wait as possible is the desirable outcome, since sighted people are used to reading instantaneously.

The most common way of measuring OCR accuracy calculates a character accuracy rate measured by the percentage of the characters that are correctly recognized. OCR developers use automated tools to evaluate this with sets of pages called test decks. The test deck pages are carefully scanned and then proofread to ensure that they are accurate. The OCR engine to be tested then recognizes the image and the resulting text is compared to the proofread text. There are many types of errors that can be evaluated (some of these overlap with each other):

- Misrecs: where the correct character is misrecognized as a different character: saying it is the letter 'c' when it is really a letter 'e':
  - Splits: where the character is turned into multiple characters, often because the contract is too light 'iii' for the letter 'm'.
  - Joins: multiple characters are incorrectly recognized as one: 'm' for 'rn'.
- Nonrecs: where the OCR engine detects the symbol, but has no idea what the symbol is.
- Drops: the OCR engine fails to put anything out for a character.
- Adds: characters are output that were not characters on the original, such as a speck of dirt coming out as a period or quote mark.

This sort of approach is quite time-consuming. Still, modern OCR systems are quite accurate, doing very well on simple text-oriented documents such as novels. If the scanning is done carefully, the accuracy can be very close to perfect, with no errors on a typical page. Highly complex documents are very difficult to read using OCR: complicated textbooks, fancy magazine articles and hand-written documents are typically unreadable. OCR users become adept at quickly judging how to get the best results out of their OCR technology, as well as recognizing its limitations.

Speed has become less of an issue over the years as computers have become faster and faster. It is usually the scanner that is the limiting factor in the speed of converting documents, especially if the pages are being turned by hand and the book is being placed on a flat-bed scanner.

Institutions scanning large numbers of books often use higher capacity production tools to deliver greater amounts of scanning, using chopping devices to remove bindings, high speed page scanners and OCR software. Much commercial OCR software is highly visual, providing detailed images of pages for direct interaction. Reading systems that integrate OCR and can be used by people who are not able to see the page are the solution for those who are visually impaired.

15.3 Reading Systems

Reading systems are designed for use by people with disabilities. They use OCR to provide the printed word in accessible form. There are three main elements of a reading system: the scanner, the OCR system and the accessible presentation.

Reading systems are also packaged differently from conventional OCR. The majority of reading systems are built on a standard personal computer, where the user is likely to also have the use of a screen reader or screen magnifier for other tasks. In addition, there are stand-alone reading systems, which bundle all of the components into a single purpose unit that reads; the user is shielded from computer capabilities of the device. The functioning of the device is the same, but the complexity is hidden to make the device less intimidating.

Once the text of the document is available, the user has many ways of accessing it. Accessible presentation technologies include text-to-speech, Braille and enlargement. These same technologies are often used to provide control feedback to users to operate the reading system. Stand-alone systems use text-to-speech almost exclusively. PC-based systems use one or more of the following access techniques.

Text-to-speech (TTS) is the most widely used technology for providing access to printed information with reading systems, because it is quite inexpensive and works for the great majority of visually impaired users. Synthetic speech sounds artificial, even after the major technology advances made over the past decade. This makes reading systems unattractive to individuals who are reluctant users of technology, especially seniors. However, the progress in speech technology means that TTS is much closer to sounding like a human narrator than the early computerized voices many consumers may still remember.

Braille technology is very popular with the segment of the blind who read Braille, as it is more precise and efficient than speech output. Because the majority of the legally blind population are seniors who generally do not learn Braille, less than 20% of the blind are effective Braille readers. However, the educational and economic success of blind persons who are Braille readers tends to be much higher than non-Braille readers. Canadian surveys in the 1990s (Campbell *et al.* 1996) showed that employment-aged Braille readers had an unemployment rate lower than the general population, and also had better economic and educational attainment.

There are certain barriers to the use of Braille; it is very expensive and not included as a standard component of reading systems. Refreshable Braille displays typically have 20, 40 or 80 characters of Braille using plastic tactile pins that pop up to display the text, emulating the feel of dots embossed onto paper as is used in hardcopy Braille. This technology is indispensable to people who are both deaf and blind, and cannot use audible speech output. Reading systems also print Braille documents using specialized printers called embossers, which punch Braille dots into paper to create hardcopy Braille books and documents.

The last accessible presentation technology is enlargement. This is the electronic equivalent of the video magnifier, but access to the underlying text using OCR can make it easier to display the text useably to the low vision reader. For example, when working in a large font, wrapping the text on a screen is easier to do with digital text compared to the same task using a picture of a page of text. The colours and contrast can be adjusted through a wider range of options than in a direct video magnifier.

Combinations of accessible technologies are quite common. Low vision users appreciate the option to both view enlarged text, as well as listen to it at the same time. This "bi-modal" display technique, originally designed for low vision users, is very popular with people with learning disabilities such as dyslexia. Braille readers often use a combination of a Braille display with TTS.

At the time of writing, the leading reading systems for the visually impaired are products from companies such as Freedom Scientific (OPENBook) and Kurzweil (the Kurzweil 1000). These are PC-based software solutions, where adding the reading system software and a scanner turn a PC into a talking reading system. Users who are not very skilled using a PC can often use these reading systems because they can be effectively operated with a couple of keys. The user places a page of print on the scanner and presses a key that tells the reading system to scan the page, do the OCR and start reading the page aloud. The other important key starts and stops the speech output reading of the page.

As an alternative to PC-based systems, Freedom Scientific also makes a stand-alone version of their reading system, called SARA. In addition, there are a variety of other reading systems produced in different countries such as the Poet by Baum in Germany and the ScannaR from Human Ware in New Zealand. These systems hide the computer details into a single housing, with the controls and speakers generally built into the front of the scanner. While these are more expensive than PC-based systems for users who already own a PC, they are far less intimidating to the non-technical user.

There is also a new portable reading machine, the Kurzweil-National Federation of the Blind Reader, which is the marriage of a digital camera to a personal digital assistant. It weighs less than 400 g (under 1 lb). The user aims the camera lens at a page of text and the Reader speaks it aloud with TTS. It is relatively expensive, costing more than a stand-alone reading machine with a flatbed scanner. The typical user would be unlikely to read an entire book using such a device, but it seems well suited to daily reading tasks like reading mail and other short documents.

Other than the K-NFB Reader, reading systems are not very portable. This is especially true when considering reading entire books. Since portability in reading is important, users have come up with a number of ways to take their reading with them. Many Braille displays are designed to be portable as a notetaker device, enabling the user to store the scanned book in the device's memory and carry around a small library of books. Given the large size of hardcopy Braille books, it is quite exciting to have hundreds of books stored inside a notetaker, which is smaller than a single Braille hardcopy volume.

There are also specialized portable devices with TTS synthesizers built in. The first type consists of TTS notetakers, either with QWERTY full keyboards or Braille chording keyboards (where text is entered by chording the six or eight Braille dots for each character). These are less expensive than the equivalent Braille notetakers which are equipped with refreshable Braille displays because of the expense of the Braille cell technology. In addition, there also devices designed to simply be text readers, such as the BookCourier and Book Port. These have built-in TTS and simple numeric pads, and read books aloud through a headphone jack.

PC users are increasingly using software that creates MP3 digital audio files using TTS, so that commonly available music players can be used to play books aloud. Reading systems such as OPENBook and K1000 have an option to create MP3 files from scanned text. A software program that performs just this task is TextAloud, which takes a text file and creates an MP3 audio file. The quality depends on which speech synthesis engine is used to create the audio: there is a significant difference in voice quality and naturalness among different speech output TTS engines.

MP3 files have become the current *de facto* standard for digital audio files. The name comes from Motion Picture Experts Group Audio Layer 3, a standard for compressing audio signals to take-up less memory space. It contrasts with the Microsoft/IBM WAV files, which are generally uncompressed and can be ten times larger for the same audio content. MP3 is a *lossy* compression technique, which means that it discards some information in the process of making the audio file smaller. However, the information that is lost is carefully chosen to be information that the human ear would generally not perceive.

MP3 file players have been incredibly popular worldwide because of their music application. The most famous device at the time of this writing is the iPod from Apple Computer (which is unfortunately not quite accessible for blind people). Because of mass production, some of these devices can be significant less expensive than similar devices custom designed for visually impaired people. MP3 players usually get their files from a PC by connecting a cable and synchronizing the device with the desired files on the PC. Since MP3 players have become so popular, these capabilities are now migrating to other devices, such as cellular phones and personal digital assistants.

To meet the wide variety of needs of different users, there are many options for utilizing text scanned in through a reading system. By examining factors such as budget, preferred access modes (Braille, large print, audio), the need for mobility and level of comfort using computers, it is possible to narrow these options down to a handful of solutions worth examining.

15.4 DAISY Technology

The DAISY standard was created by a technical consortium of organizations that provide accessible books to people with disabilities. DAISY stands for Digital Accessible Information SYstem. The consortium was created by the leading libraries for the disabled from around the world to develop a digital successor to audio books on tape. The goal of accessible books is to provide an alternative way for a visually impaired person to have the same or better access to a book that a sighted person has. This includes more than just accessing the content. A sighted reader can browse through the printed book, skim the text, or look at the table of contents or index and go immediately to the information they are seeking. Some people choose to read the ending of a book before the rest of it. These are challenges that the DAISY standard is meeting.

The vision of the DAISY Consortium is: "that all published information is available to people with print disabilities, at the same time and at no greater cost, in an accessible, feature-rich, navigable format". Unlike reading systems, which focus on providing users themselves with the ability to turn print books into accessible form, the member libraries of the DAISY Consortium specialize in delivering accessible books directly to users.

The current state of the art in much of the world (including the United States) for talking books for visually impaired people is 4-track audio cassettes. Compared to the technology used by rest of the population, this is decades behind the times. The typical book requires multiple, bulky cassettes and a specialized player to listen to the cassettes. Going to a specific page of the book is difficult and time consuming. The table of contents and index from the printed book are also impractical to use. The world needed to move beyond the analogue audio cassette into the digital world, and that is the need DAISY was created to address.

One important goal in developing a new talking book standard is to build critical mass around a specific format. That way, many different parties can work on different aspects of the problem of book accessibility with confidence that their work will be interchangeable. For example, once there was a CD-audio standard, recording companies could release music albums on CD and know that customers could put that music CD into audio-CD player and music would come out of the speakers. This is exactly what the DAISY Consortium was aiming for: a new standard where technology makers could build players and accessible book publishers could publish books and know that the books would play in the players. It also meant that books created by one library could be used by patrons of other libraries.

The DAISY standard describes a format for digital accessible books. It is a flexible format designed to allow for many needs and methods of providing accessible books. As with many formats, the term DAISY is used in many ways. A DAISY book refers to a digital book in the DAISY format. A DAISY player is an appliance or a piece of software that can play a DAISY book. DAISY is similar to the MP3 format. Just as songs can be placed in the MP3 format, delivered to users by CD or downloading, and played on an MP3 player or on a PC running MP3 playing software, DAISY books can be used the same way.

DAISY pulls three types of information together: the content, the structure and the linkages between the content and the structure. The content can be one or more of the following: text, recorded audio or images. The structure information is modelled after the structure used in print books. A table of contents might contain chapter titles and page numbers. The chapters and pages would be the structure. The linkages make it possible to navigate the book as a sighted person would. An example will illustrate this.

The most typical DAISY book today is an audio book with chapter structure. A sighted person can look at the table of contents of the print version of the book, choose chapter six and go immediately to the audio narrative of that page. With a DAISY version of the book, the user can listen to the table of contents, and link directly to chapter six and begin hearing chapter six. If a teacher directs students to go to page 152, the student with a disability should be able to go directly there. This is the essence of what DAISY aims to deliver: the same level of access to content that people without disabilities enjoy.

Because DAISY is working with different kinds of content, it builds on standards already in place. For example, the audio content of a DAISY book is often provided in an MP3 format. As a result, inside a DAISY format book you might find that each chapter is a separate MP3 file. The DAISY player is designed to handle this and avoid the necessity for the user to learn the inner workings of the reading system.

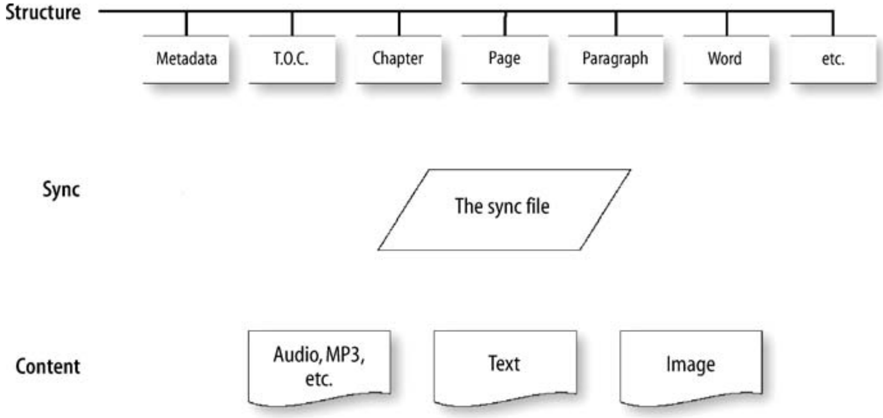


Figure 15.4. DAISY book structure

The text and structure of a DAISY book are in an eXtensible Markup Language (XML) format. Like many formats, DAISY comes in different versions and flavours. An encouraging trend is the unification of the latest DAISY format with the ANSI/NISO Z39.86-2002 standard, so that the international DAISY organizations will be using the same standard as the National Library Service for the Blind and Physically Handicapped of the U.S. Library of Congress. The National Instructional Materials Accessibility Standard (NIMAS), adopted by the U.S. Department of Education for textbook accessibility for K-12 students, is explicitly defined as a subset of the ANSI/NISO standard.

Because the DAISY standard is a format, it is not tied to a specific tangible medium (see Figure 15.4 for the DAISY book structure). It is possible to deliver DAISY books on most digital media, such as CDs, DVDs, memory cards or sticks, or simply make them available for download over a network such as the Internet.

Although there are six different types of DAISY books, at present there are two main types being distributed: full audio and full text. Most players can only play one type or the other, although this is likely to change in the future.

15.4.1 DAISY Full Audio Books

The first type is the audio DAISY book, where the core content is a recording of a human narrating the book. Depending on the book, you may have different amounts of structure information. A single essay may be one page and have no navigation. There is one piece of audio, and the reader listens to it. Usually, the books will have chapter information. On newly produced books, it is also common to have page number information. This information is what makes it possible to skip to the desired content, unlike an audio tape where it is necessary to fast forward through everything between the current and the desired locations.

The actual format of the audio content inside a DAISY book is based on existing audio standards, such as MP3 or WAV. You can think of DAISY as a wrapper that goes around the book content and makes it easy to navigate, but at the core level you will see text, audio and graphics files in the common industry standards, such as XML text, MP3 and JPEG. Since an accessible talking book can contain all of these different types of information, it made sense to focus the DAISY efforts on adding the navigation elements and using the main existing standard formats for text, audio and graphics.

There are countries, such as Denmark and Japan, where the main libraries for the blind have shifted all new audio books for the disabled to the DAISY format. Many other organizations are in the process of moving to this new format. Current audio DAISY books are primarily provided on CD-ROM media because of the large size of the audio files. Plans are already being made to distribute these books in other digital formats, such as solid-state memory cards or memory sticks.

15.4.2 DAISY Full Text Books

The second main type is the full text DAISY book, where the content is the digital text of the book, much as it would be if it were a webpage or a word processor file. There is no human narration, which makes this kind of book much less expensive to produce. To access the content, visually impaired users read it using a voice synthesizer or a Braille display, much as they do on reading machines. In fact, the majority of these text DAISY books available today are initially created by people using reading machines to scan books.

The text XML used in DAISY looks quite similar to HyperText Markup Language (HTML), the format that the World Wide Web uses. For example, a paragraph of text on the Web would have a paragraph tag around it, which looks like

<p> The paragraph goes here. </p>

DAISY XML uses the same kind of tag for paragraphs. It adds additional tags that are useful for navigating books, such as page number tags, chapter tags, footnote tags and so on. By adding these tags into the text content, the DAISY player can support skipping forward by paragraph or page or chapter, going to a specific page or chapter, skipping all footnotes and so on. But, the information has to be inside the XML to be able to use it.

Since it was designed for web pages, HTML is not well suited for book access, and thus lacks support for these important navigation features. If you were to bring a book into a web browser such as Internet Explorer or FireFox, it would show up as one giant, long page. DAISY uses additional tags to provide the page and chapter navigation missing from HTML.

Bookshare.org in the United States is an example of a text DAISY library. The books are primarily scanned by people with disabilities, teachers and volunteers. Other books are provided directly by publishers and authors in a digital text format and converted to DAISY. Because of the relatively small size of DAISY text files (less than 1 MB per book), these books are easy to download quickly over the Internet, even on slow connections.

NIMAS is a text-only DAISY book format with a few minor additions (a few extra metadata tags, such as grade-level). As of 2007, the United States has a national repository of digital textbooks for primary and secondary students in the NIMAS format. One of the goals of settling on a single format was to make life easier for publishers, especially those required by law to provide accessible versions of their textbooks. It is also easy to turn NIMAS format books into other accessible formats, such as Braille.

Table 15.1 shows how the beginning of a public domain book, *Wuthering Heights* by Emily Bronte, would appear in three different text formats: ASCII, HTML and DAISY XML. You can see how each format builds on the previous one, adding helpful information. This is a particularly simple version of the DAISY book, but it shows the support for metadata and navigation tags that make the DAISY book more useful to a user.

15.4.3 DAISY and Other Formats

The electronic formats used for books and other documents are growing together. For example, it is very easy to convert from one XML format to another. Converting from the commercial ebook XML formats, such as DocBook and the Open eBook, into DAISY is very easy because they are constructed using similar concepts.

Text DAISY can also be used to generate Braille formats and documents. Braille translators such as the Duxbury program can take DAISY text XML and turn it into a form of Braille shorthand called grade II Braille. Duxbury uses the structure and navigation information in the DAISY file to better format the Braille. For

```
             ,WU!R+ ,HEI<TS           #A

…
  ,2G9 ,3T5T
  ,WU!R+ ,HEI<TS                      #D
  0,EMILY ,BRONTE
  ,,*APT] ,I
  #AHJA4 -- ,I H J RETURN$ F A VISIT 6MY
L&"L -- ! SOLIT>Y NEI<B\R T ,I % 2 TR\#D
)4 ,? IS C]TA9LY A B1UTI;L C.TRY6 ,9 ALL
,5GL&1 ,I D N 2LIEVE T ,I CD H FIX$ ON A
SITU,N S -PLETELY REMOV$ F ! /IR (
SOCIETY4 ,A P]FECT MISAN?ROPI/'S H1V53 &
,MR4 ,H1?CLIFF & ,I >E S* A SUITA# PAIR
6DIVIDE ! DESOL,N 2T U4 ,A CAPITAL
FELL[6 ,HE LL IMAG9$ H[ MY HE>T W>M$
T[>DS HM :5 ,I 2HELD 8 BLACK EYES )DRAW
S SUSPICI\SLY "U _! BR[S1 Z ,I RODE UP1
& :5 8 F+]S %ELT]$ !MVS1 )A J1L\S
RESOLU;N1 / FUR!R 9 8 WAI/COAT1 Z ,I
ANN\NC$ MY "N4
  ,8,MR4 ,H1?CLIFF80' ,I SD4
  ,A NOD 0 ! ANSW]4
```

Figure 15.5. Grade II Braille example

Table 15.1. Format examples

ASCII version	HTML version	DAISY version
Begin Content Wuthering Heights by Emily Bronte CHAPTER I 1801. – I have just returned from a visit to my landlord – the solitary neighbour that I shall be troubled with. This is certainly a beautiful country! In all England, I do not believe that I could have fixed on a situation so completely removed from the stir of society. A perfect misanthropist's heaven: and Mr. Heathcliff and I are such a suitable pair to divide the desolation between us. A capital fellow! He little imagined how my heart warmed towards him when I beheld his black eyes withdraw so suspiciously under their brows, as I rode up, and when his fingers sheltered themselves, with a jealous resolution, still further in his waistcoat, as I announced my name. 'Mr. Heathcliff?' I said. A nod was the answer.	<html> <body> Wuthering Heights Begin Content Wuthering Heights by Emily Bronte CHAPTER I 1801. – I have just returned from a visit to my landlord – the solitary neighbour that I shall be troubled with. This is certainly a beautiful country! In all England, I do not believe that I could have fixed on a situation so completely removed from the stir of society. A perfect misanthropist's heaven: and Mr. Heathcliff and I are such a suitable pair to divide the desolation between us. A capital fellow! He little imagined how my heart warmed towards him when I beheld his black eyes withdraw so suspiciously under their brows, as I rode up, and when his fingers sheltered themselves, with a jealous resolution, still further in his waistcoat, as I announced my name. 'Mr. Heathcliff?' I said. A nod was the answer.	<?xml version='1.0' encoding='ISO-8859-1' standalone='yes' ?> <?xml-stylesheet href="RevStd.css" type="text/css"?> <dtbook3> <head> <title>Wuthering Heights</title> <author>Emily Bronte</author> < /head> <book lang='en'> <frontmatter> <doctitle>Wuthering Heights</doctitle> <level1 class='Intro'>Converted From ASCII and repaginated</level1> < /frontmatter> <bodymatter> <level1 class='part' id='start'> <h1 class='part'>Wuthering Heights</h1> p id="para_000010">Begin Content</p> <p id="para_000011"></p> <pagenum id="page_000001" page='normal'>1</pagenum> <p id="para_000012"></p> <p id="para_000013"></p> <p id="para_000014"></p> <p id="para_000015">Wuthering Heights</p> <p id="para_000016">by Emily Bronte</p> <p id="para_000017"></p> <p id="para_000018">CHAPTER I</p> <p id="para_000019"></p> <p id="para_000020">1801. – I have just returned from a visit to my landlord – the solitary neighbour that I shall be troubled with. This is certainly a beautiful country! In all England, I do not believe that I could have fixed on a situation so completely removed from the stir of society. A perfect misanthropist's heaven: and Mr. Heathcliff and I are such a suitable pair to divide the desolation between us. A capital fellow! He little imagined how my heart warmed towards him when I beheld his black eyes withdraw so suspiciously under their brows, as I rode up, and when his fingers sheltered themselves, with a jealous resolution, still further in his waistcoat, as I announced my name.</p> <p id="para_000021">'Mr. Heathcliff?' I said.</p> <p id="para_000022">A nod was the answer.</p>

example, the Bookshare.org service uses tools to create XML from scanned books in formats such as Microsoft Word or Rich Text Format and then automatically creates a Braille version from the XML. Users have the choice of either DAISY or BRF, as the Braille format is known. The text shown in Table 15.1 is presented in Grade II Braille in Figure 15.5. As noted before, it is also easy to create audio files, such as MP3 format files, from text DAISY ebooks.

15.5 Players

All the effort to create accessible books becomes worthwhile when a user sits down to read. Since audio is the most widely used form of these books, the devices that provide this capability are typically called players. Players come both in physical form, like a cassette player, and software applications that run on a more general purpose device, such as a personal computer. The main considerations in choosing players are security, navigation, durability and portability.

Security has two elements: specialized formats and digital rights management ("DRM"). Both security elements derive from the requirement to prevent accessible books for the disabled community from leaking into the general population. These books are often created without royalty obligations under law or license; the resulting social bargain with publishers and authors is to provide access to people with disabilities without discouraging the sale of commercial printed books. When the specialized formats were hardcopy Braille books and 4-track audio tapes that could not be played on mass market cassette players, the security requirement was addressed. With the new purely digital formats, new forms of protection are being devised to respond to publisher and author concerns about possible misuse of the content. These DRM solutions intend to provide easy access to the desired users, while discouraging access to those who are not qualified.

Different providers of ebooks implement different digital rights management programs to restrict access to authorized individuals and organizations. One of problems of early DRM approaches for commercial electronic books was that they tended to prevent access by the adaptive technology used by people with visual disabilities. This was ironic, since electronic books are fundamentally more accessible than print books! This problem was discussed at length in a white paper entitled "*The Soundproof Book*" by Kerscher and Fruchterman (2002).

The DAISY standard does not define a specific kind of DRM, which leads to variations, but all of these variations are fundamentally accessible. Two examples will demonstrate this.

Bookshare.org has a seven point DRM plan that it has implemented and reviewed with publishers and authors. The key points are:

- Qualified users: users have to provide certification that they have a *bona fide* print disability.
- Contractual agreement: every person who downloads books from Bookshare.org has to agree to not provide them to third parties (or post them on the Internet).

- Copyright notice: each copyrighted book has an extensive notice that points out that it is only intended for people with print disabilities, as well as noting who the copyright owner is.
- Encryption: each book is encrypted with the individual user's password.
- Watermarking and fingerprinting: each book is marked with a digital watermark that identifies it as coming from Bookshare.org, as well as a fingerprint identifying the specific user who downloaded the book.
- Security database: Bookshare.org keeps a database of who has downloaded each book.
- Security watch program: unusual downloading patterns are watched for possible abuse.

Together, these DRM policies discourage abuse of the copyrighted materials. However, since users read text so many different ways, the system does permit users to decrypt the books and load them into different devices, create Braille, enlarge the text, convert it to MP3 audio and the like. The main protection against abuse is actually social norming: someone who abuses the system threatens access for the entire community. Community solidarity and peer pressure are reasonably effective in encouraging responsible behaviour.

Recording for the Blind & Dyslexic is providing DAISY full audio books at the time of this writing and does not allow the full decryption of books. Users are required to provide proof of disability and sign agreements, but DAISY books are not customized to the individual. They are shipped out on CDs. To play a DAISY audio book, you need to enter a PIN (personal identification number) into the DAISY player, whether it's a separate hardware player like the Victor Reader or a software player running on a PC.

Both of these approaches have worked fairly well. Users would prefer no DRM requirements, but are willing to put up with these to gain access to books.

Navigation, the second key player consideration, is what makes a DAISY format book much more usable. The tape machines were a great advance for their day, but fast forwarding through audio tapes is unwieldy and frustrating to today's user. To explore navigation, try to imagine how a blind person would perform the navigation tasks that a sighted user can do easily:

- Going to a specific chapter or page.
- Skimming.
- Looking up a term in the index and then going to where it is used.
- Accessing a glossary of terms.
- Reading a footnote.
- Accessing a chart, table, picture or graphic.

A simple user interface is also an important part of player design. It is not helpful to have a capability that few users can understand and use.

Durability is especially important for large government libraries that provide free or discounted hardware players. Expecting a player to have a useful life of ten years under hard use is a very different requirement than what is expected of typical consumer electronics.

Portability is increasingly a top consideration in player choice. Print books are quite portable, and people read them in every conceivable location: in bed, in a comfortable chair, on the bus, on the beach and so on. Users of accessible books are trying to achieve the same flexibility. Active adults and students require portability.

There are a growing number of specialized devices and software applications that meet the need for players. These devices often focus on portability, since the software applications are generally tied to the use of a personal computer.

Four-track audio tape players are in wide use today. These range from heavy duty players provided at no cost by a government agency (such as by the Library of Congress in the U.S.) to modified commercial cassette players. As noted before, almost all libraries for the blind have either discontinued the use of these players or are actively planning for making them obsolete. However, the widespread use of these players means that blind people will be using them for many years to come, even as the new technologies are deployed.

DAISY audio book players also come in two major flavours: heavy duty desktop models, provided by national libraries (such as the RNIB in the U.K.) and more portable versions based on Walkman-style music players (see Figure 15.6). At present, these players work with DAISY audio books delivered on CDs. In addition to the traditional audio controls of play, stop, forward and back, these players all have additional controls to take advantage of the DAISY navigation capabilities.

Devices that can act as players for text-based books include built-in access technology. For example, the BookCourier and BookPort products can accept text downloaded from a computer and play it through an earphone using text-to-speech technology. Blind people also use a range of notetakers with text-to-speech capabilities, such as the PAC Mate and VoiceNote products. Notetakers are more

a b

Figure 15.6. Two photos of HumanWare DAISY hardware readers

than just players: they also have Braille or QWERTY-style keyboards that support text entry and include a range of built-in applications.

Notetakers with built-in Braille displays have revolutionized access to Braille. Blind Braille readers are thrilled with the ability to carry an entire library of electronic books in their notetakers. Hardcopy embossed Braille is quite bulky, so a single small memory card can hold books which would consume many bookshelves in paper form. This also greatly reduces the cost of providing Braille, once the investment is made in a notetaker with a display. For example, a typical digital Braille book is 400 kB. An inexpensive 512MB Compact Flash storage card can hold more than one thousand books of this size, and this kind of card can be inserted into the major notetakers currently in use. With the rapid increases in memory size at constant price, personal libraries of digital Braille books are likely to grow rapidly.

In addition to the specialized players designed for the visually impaired community, there are many commercial devices that can be used. Typically, the user interfaces present challenges because the blind were not part of the design process. Still, the low cost of these devices make them quite attractive, especially for people who cannot afford the specialized equipment and do not have a government subsidy to help acquire it. To utilize these commercial devices takes more effort and is not yet suitable for the average user.

For example, commercial music players can play book content in a format such as MP3. If the DRM permits it and if the format is suitable, audio books can be downloaded into these players. As was discussed earlier, many software applications can turn text into MP3 with TTS.

In addition, Audible.com is a commercial service that offers digital audio books in its own proprietary format. They have both a software application to play their books on a PC, as well as a custom hardware player that is similar to an MP3 player. This service is quite popular with blind people who can afford to subscribe to it. Furthermore, it is available globally, and Audible.com has expanded its collections of audio books into languages other than English (specifically French and German).

As accessible books become more available, and as the kinds of players proliferate, we should expect to see a great deal more consumer choice about access. Lowered costs and easy to use players mean that accessible books are likely to reach a much larger proportion of people with visual disabilities around the globe. It is not hard to imagine a cellular phone soon having the capability to store and play an accessible book either as recorded audio or as synthetic speech.

15.6 Accessing Textbooks

Textbooks are growing in importance as a special case of book access. Many countries promise access to education to people with disabilities, and that requires delivering accessible educational content. The quality requirement for textbooks is very high, in order to ensure the disabled student has access comparable to that available to the student without a disability. New efforts are mandating that textbooks come directly from publishers in an accessible digital form.

The complexity of modern textbooks makes access more challenging. Textbooks often have complicated layouts, with sidebars and other fancy graphical elements, making scanning with OCR difficult. Individual school districts have spent considerable money making textbooks accessible. It is not uncommon for a US school to spend US $ 10,000 or more to convert a single advanced math textbook into Braille. Math and science symbols are beyond the current state of the art for OCR, as well. Finally, the economics of the textbook industry, especially at the post-secondary level, drive rapid revisions of textbooks to encourage the sale of new textbooks and discourage the sale of used textbooks.

There are organizations that focus a significant proportion of their effort on the needs of students. For example, Recording for the Blind & Dyslexic has a network of volunteers that specialize in making textbooks accessible by recording them aloud. These volunteers are typically drawn from the fields covered by specific textbooks, and can verbally describe graphs, charts, formulas and other challenging document elements. This skill is well beyond the capabilities of automated approaches such as OCR.

The DAISY format was designed with the special needs of textbook access in mind. In addition, there is growing pressure on publishers to provide digital versions of textbooks to make accessibility more feasible. Many jurisdictions have passed laws mandating publishers provide digital versions of books to students with disabilities. The United States Department of Education just announced its support for the National Instructional Materials Accessibility Standard, which is based on the DAISY and ANSI/NISO Z39.86 standards, and will require all publishers of textbooks at the primary and secondary levels to produce digital versions in this format to meet accessibility needs. One goal of this standardization effort is to make it easier and less expensive to produce a high quality textbook, through standardizing the format that publishers can use for export. For example, it is quite easy to take digital books in the DAISY format and create a Braille or digital audio book with good navigation capabilities. Starting with the original, publisher quality text, removes the need to OCR textbooks and significantly reduces the barrier to access for textbooks.

15.7 Accessing Newspapers

Newspapers pose a unique accessibility challenge. They are not practical to scan using a reading system, and traditional means of producing accessible documents are impractical for information provided daily (and which is most useful on the day of publication). Many newspapers are available on the Web, but each newspaper approaches web publishing differently.

One mechanism used in several countries is the radio reading service. Volunteers read the newspaper aloud, and the recording is broadcast over a specialized radio service. Users often receive specially tuned radio receivers for listening to these broadcasts, which cannot be heard on typical AM/FM receivers. These services are generally free. However, they cannot read an entire newspaper, and the listener has no direct control over the content chosen or the order of reading.

The Talking Newspaper Association of the UK creates audio tapes every week for its subscribers, with subscribers able to choose which newspapers or magazines to receive. Recently, the service was expanded to offer both audio CDs and digital text files.

A popular service in the United States is the NFB-Newsline® service, operated by the National Federation of the Blind. Every morning, the Newsline servers receive the daily editions of many different newspapers. Users telephone in and listen to the newspapers and articles they select, with the information delivered by text-to-speech over the telephone. Newsline is now piloting an initiative to provide the same content in Braille and DAISY formats through the Bookshare.org service, so that the papers can be downloaded into PCs and notetakers. Newspapers are highly structured and can easily be converted into the DAISY format, making navigation easy.

Many of the same approaches are being tried with magazines. All of these approaches to periodicals share the basic goal of accessible documents: to provide at least the same quality of access to the printed word as that enjoyed by the general public.

15.8 Future Technology Developments

Technology serving the blind is aligning more closely with mass market technologies which are improving with relentless speed. Conversely, the affordability gap between custom technology for the visually impaired and consumer technology will remain and may widen. Therefore, reading technology is likely to evolve rapidly on the back of the main computer, telecommunications and portable device markets. This overlap between the interests of people with disabilities and those without disabilities will create expanded and improved reading options. This vision of common interest has led to the universal design movement, which is playing an increasing role in accessibility regulation and design. In many cases, improvements aimed at assisting people with disabilities are also found to aid those without disabilities. This is a compelling reason to support universal design. For example, street curbs with sloping ramps at intersections were introduced to help people in wheelchairs, but these same ramps help far more people on bicycles or people pushing baby strollers and shopping carts.

The steady increases in processing power, storage, portability and device connectivity will show up in several ways. These changes can be illustrated by imagining a future generation standard cell phone, which can be used in myriad ways by visually impaired users through simply adding software. Its built-in camera is capable of reading documents and text in the environment, providing the functionality of a reading system. It can tap a huge library of books or periodicals over the network with a spoken command. These materials can be read with a human sounding voice synthesizer, or transferred seamlessly to any nearby visual or Braille display. The built-in satellite location chip begins to address the issue of personal mobility. The user does not know or care whether the processing or storage is occurring on the cell phone or the network: it simply works seamlessly. All of these capabilities are useful to people without disabilities, and so the cost for these capabilities will be within reach of even the poor.

The ultimate DAISY book will become more practical and cost-effective as it contains all sorts of exciting content. Imagine a textbook that is fully accessible to all people. All of the text in the book could be provided in audio form narrated by a human (or equally, a human sounding TTS) and in digital text form, which can be accessed using a visual or Braille display or voice synthesizer. The illustrations and graphics could have both the colour images and a detailed description of the illustration in text and/or audio. Every table of contents entry and index entry could be hyperlinked so that a reader could jump to that item or hear a dictionary definition of a given word. DAISY is powerful enough to have all of these different kinds of content in a single book, and to link them all together seamlessly.

Today, this kind of book would be very expensive to create. As time goes on, many of the costs of producing this ultimate book will decrease. Furthermore, it will appeal to the mass market. Recognition of this has led the DAISY Consortium to start positioning the DAISY format as being the ideal electronic book format for everybody, not just people with print disabilities that interfere with the reading of printed books. As electronic books become a commonplace alternative to print books, people with disabilities will benefit, as long as they are not locked out by poor design decisions.

The pace of each of these technical innovations is hard to predict precisely, and not all of them may come to pass in the next decade. However, the likelihood that they all will happen is high. The outcome of the confluence of interest between the general consumer market and people with visual disabilities, especially in the form of the cell phone of the future, will lead to expanded independent access to the tools of literacy and content.

15.9 Chapter Summary and Conclusion

15.9.1 Chapter Summary

Most computer-based access technologies use a digital text file as a key element in the system. The best way to get this digital text is directly from the publisher, but this is logistically difficult in many cases. Since much information still resides as print as in books, letters, bills, bank statements, newspapers and magazines, how is this information to be made accessible to the visually impaired? The answer is to use optical character recognition (OCR) technology to create an electronic file that can then be turned into accessible Braille, large print or audio.

In this chapter, the fundamentals of OCR technology were first described. Different reading systems built with OCR were presented. Explored next was the new international standard for digital talking books, the DAISY standard, and its associated players. Methods of accessing textbooks and newspapers were discussed next, highlighting new efforts to obtain these directly from publishers and avoiding the need to OCR these challenging documents. A discussion of future technological development closed the chapter.

15.9.2 Conclusion

The goal of parity in access to print for people with visual impairments is coming within reach. The power to choose any book or document, and the decisions of how and where to read it, will continue to shift into the hands of individual readers, without the need for human intermediaries. The DAISY Consortium's vision will shift from a Utopian concept to a practical expectation: "*that all published information is available to people with print disabilities, at the same time and at no greater cost, in an accessible, feature-rich, navigable format*". Independent access to literacy will lead to greater educational, employment and personal opportunities for all people with vision impairments around the globe.

Questions

Q.1 Examine the economics of accessible book production. Consider factors such as book acquisition, converting the book into accessible formats such as audio, Braille and digital text and quality control. Can you determine ways to make access more cost-effective?

Q.2 Examine the methods and economics of accessible book distribution. What will replace the traditional approach of providing audio tapes and Braille hardcopy books through a free postal system? Will different media continue to be shipped, or will an electronic distribution system become the dominant mode? What are the advantages of different possible methods?

Q.3 How well do you think automated translation of electronic texts works to and from languages (*i.e.*, English to Spanish and *vice versa*) you understand well? Try out some services that are available for free on the Web. Is this appropriate for solving the access problem, say for non-English speakers?

Q.4 Developing countries generally cannot afford extensive investment in accessible book programs at present. Choose a country of interest and find out what that country is doing today for accessible books. Learn about the books and materials they are focusing on. For example, many choose to concentrate on the national curriculum and examinations, since these have a significant impact on economic opportunity. Does your country have local languages where there is demand for accessible content? For your chosen country, try to learn or project what the next five or ten years will hold.

Q.5 Investigate how copyright laws make it easier or harder to share accessible books across borders. What are two or three changes you would recommend to improve this situation?

Projects

The following projects are intended to provide direct insight into the accessibility process.

P.1 Download a DAISY text ebook and look at the XML text file in a simple text editor. You will be surprised at how simple the tags are that make DAISY files more accessible.

P.2 Download a Braille electronic book in the grade II compressed format. Decode two paragraphs of Braille text using contractions tables, such as those at http://www.brailleauthority.org.

P.3 How would you take a standard cell phone and make it into a book player for the blind, or using a cell phone with a camera as a reading machine. How well would that work, and what would the biggest challenges be?

P.4 If you have the interest and commitment to stick with it, consider volunteering to read to a blind person an hour a week. This will provide insight into the access requirements that players need to meet.

To find digital ebooks that are freely downloadable, go to http://www.bookshare.org and search on an author from the nineteenth century or earlier (Shakespeare, Twain, Dickens, *etc.*) and choose one of their books. These books are public domain and available to anybody in the world.

References

Campbell, N., Sanders, J., Losier, J.P., and Owen, V., 1996, Literacy and technology, A Canadian perspective, The World Blind Union Forum on Literacy, March 26

Kerscher, G., and Fruchterman, J., 2002, The Soundproof Book: Exploration of Rights Conflict and Access to Commercial EBooks for People with Disabilities, FirstMonday

Resources

Library resources for the print disabled

DAISY Consortium http://www.daisy.org/
National Library Service for the Blind and Physically Handicapped (NLS) of the Library of Congress http://www.loc.gov/nls/

Recording for the Blind & Dyslexic (RFB&D) http://www.rfbd.org/

Bookshare.org http://www.bookshare.org/
Royal National Institute of the Blind (RNIB) http://www.rnib.org.uk/
The Swedish Library of Talking Books and Braille (TPB) http://www.tpb.se/
Japanese Society for Rehabilitation of Persons with Disabilities http://www.dinf.ne.jp/
American Printing House for the Blind (APH) http://www.aph.org/
The Force Foundation http://www.force-foundation.org.uk/

Adaptive technology products

Freedom Scientific http://www.freedomscientific.com/
HumanWare (HumanWare) http://www.humanware.com/
Kurzweil Educational Systems http://www.kurzweiledu.com/
Plextor http://www.plextalk.com/in/
Telex http://www.telex.com/Talkingbook

OCR products

ScanSoft http://www.nuance.com/scansoft/
ABBYY http://www.abbyy.com/

Formats and technology standards (be sure to check for updated versions)

DAISY version 2.02 http://www.daisy.org/publications/specifications/daisy_202.html
ANSI/NISO Z39.86-2002 http://www.niso.org/standards/resources/Z39-86-2002.html
NIMAS http://nimas.cast.org/
XML http://www.w3.org/XML/Core/#Publications
Braille http://www.brailleauthority.org

Audio books

Audio Publishers Association http://www.audiopub.org/
Audible.com http://www.audible.com/

Newspaper links

International Association of Audio Information Services http://www.iaais.org/
Talking Newspaper Association of the UK (TNAUK) http://www.tnauk.org.uk/
NFB-Newsline® http://www.nfb.org/Newspapers_by_Phone.asp

16 Designing Accessible Music Software for Print Impaired People

Learning Objectives

The provision of information in alternative formats is essential for people with print impairments and this task has traditionally been performed by specialist organizations. However, this approach is normally implemented as an afterthought and is rarely integrated with mainstream solutions. These independent applications are then at a disadvantage whenever software versions or operating systems are updated. In order to make this integration process easier, and provide more intuitive designs for the future, it is essential that Design for All and accessible design methodologies are more widely adopted. The use of agreed standards is another way in which greater integration can be achieved, and this can be seen in the adoption of standards for web accessibility.

The provision of music in alternative formats is one area where only partial progress has been made, but this reflects the general fragmentation in music encoding formats. Over the last ten years many different formats have been suggested but no single agreed format has emerged. It is therefore very difficult to design software that can take advantage of advances in multimedia environments when no general agreement has been established for music encoding within these environments. However, it is possible partially to overcome this problem by adhering to the sound precepts of Design For All and by adopting thoughtful approaches to object-oriented software design.

For such approaches to succeed there is a parallel need for a wider understanding of accessibility concepts, which are often perceived as peripheral considerations. In fact, the definition of accessibility is rapidly changing and many designers are now including accessibility alongside usability.

This chapter provides an introduction to accessible information processing and focuses on the provision of music to print impaired people. From this chapter you should have an initial understanding of:

- The provision of music notation for print impaired people.
- The key tenets of Braille music.
- The creation of music notation within multimedia environments.
- The scope of the Design For All paradigm.

- The need to consider *accessible* information processing from the very beginning.
- The importance of relevant standards.

16.1 Introduction

As music scores are no longer limited to a sheet of printed paper, it becomes possible to reformat musical information in such a way that access to these scores can be provided in multimedia formats. These new ways of making music more accessible provide new opportunities for creating, modifying and remodelling the musical content of digital files for artistic or scientific purposes. Consequently, this offers new opportunities within educational and vocational environments to those who are unable to read music scores.

This chapter will give an overview of the methods used to make music accessible for people with a print impairment. These methods provide an enhanced level of accessibility to music scores and allow visually impaired musicians to model and explore musical content for their own purposes.

16.1.1 Print Impairments

First, the target audience for accessible music software must be identified. The term *print impaired* refers to anyone who has a problem understanding the traditional print version of information. In the musical domain, this refers to anyone who cannot read a music score. The solutions discussed in this chapter often refer to solutions which were primarily conceived for blind users, but many of the solutions have proved valuable for larger groups, including the partially sighted and the dyslexic, as the problems addressed are often very similar. This shows the benefit of well-defined user requirements, where simple transformations of the primary development goals can be rewarding for a wider body of users.

The breadth of print impaired users is similar to that for general users, and the target audience for accessible music software also includes those who provide accessible music in different environments, as illustrated in Table 16.1.

In most cases, visually impaired people can have access to visually coded information by means of touch (tactile), audio, haptics, print enlargement, or by a combination of these means. These media must be considered as basic communication channels. In other words, it is not sufficient to adopt simple transposition procedures to render visual information into non-visual information. It is useful to consider an analogy from language translation: if this is done just one word at a time into another language much of the meaning is lost. The same is true for translations into accessible formats.

For most print impaired users, the main point of access to information is the written word. This is usually provided in the form of paper-based accessible formats, audio formats, or through hardware connected to a computer screen. Print impaired people are entitled to read the same material as their fellow citizens, at the same time and at no additional cost to the individual in order to avoid social

Table 16.1. Overview of the user groups involved in accessible music

Individuals	Education	Industry
Visually impaired people	Music schools	Transcription centres
Print impaired people	Schools for blind and visually impaired people	Music publishers
Visually impaired musicians	Music libraries	Music content providers
Print impaired musicians	Libraries for the blind	Music software developers
	Conservatories	Organisations developing and distributing tools and aids for visually impaired people

exclusion. Naturally, creators and those who add value to creative work, have legitimate economic and moral interests that should be respected. However, while there is a commercial market for a limited range of 'accessible' materials, most of these materials have to be created by specialist agencies operating on charitable funds or social subventions. This means that, in practice, only a small proportion of the material currently published becomes available in accessible formats.

16.1.2 Music Notation

Music is artistic content, and for many people is a crucial factor in their development of cognitive dimensions, be they expressive, social, educational or creative. The visually impaired individual should, therefore, have the opportunity to access both *performed music* and *represented music*. Performed music refers to audible music, and includes all forms of real music, both traditional and innovative. The term *represented music* is any kind of coded music, including Braille music, computer-accessible music codes, or talking music.

Many visually impaired people can access *performed music*: that is, they can play, sing, navigate through CD-ROMs or the Internet (as with spoken language); but very few have access to written music and the consequent step towards musical literacy in the wider sense.

The main provision for print impaired musicians accessing music notation over the last century has been Braille music (see Section 16.2.1.1). The increasing use of computers as a medium for creating and using music notation (in both domestic and commercial environments) opens up new and interesting possibilities for music coding and representation. More importantly, these developing initiatives for incorporating more intelligence into musical representations open up new possibilities for addressing the needs of print impaired users. In the past many people have by-passed music notation and become successful performers, creators and users of music without the need for music notation. This is especially true among dyslexic musicians. Work in this field also opens up the music score to make it available to 'niche' markets and increases the scope of the mainstream music products currently on the market.

16.2 Overview of Accessible Music

This section gives an overview of the instruments and tools that are available for producing accessible music and the methods that are currently used.

From a production logistics point of view, the accessible solutions associated with providing for print impaired people present several problems. The accessible market is a niche market, so the problem becomes both a logistic problem and a design problem. Solutions are primarily designed with the mainstream market in mind. Once these user requirements have been adequately met, secondary user needs can be considered. However, as the accessibility requirements have been considered as a secondary solution or an afterthought, the solution at this stage is very often a *piggyback* solution. This creates a very poor design environment and fails to incorporate the basic ideas of Design and Accessibility For All (http://www.design-for-all.org/). The original software is usually designed with very robust and modern design methodologies, yet a quick solution is designed for the niche markets.

One of the most important requirements of musical presentation in Braille or Talking Music formats for print impaired users is the ease of comprehension of the context of the musical material. This means that the important information regarding context deduction must be clearly perceivable. That is, the end-user must be able to establish their location in the musical material and so be able to understand the functions of the musical material at hand. The end-user must not be overloaded with redundant information, but at the same time be provided with enough musical material to be able to deduce its function.

16.2.1 Formats

16.2.1.1 Braille Music

Ever since Louis Braille invented his system for representing music, blind musicians have been able to obtain scores in the Braille music format. There are international guidelines for Braille music notation (Krolick 1996). Materials in Braille music make up the largest portion of the available alternative formats, and include the standard repertoire for most instruments, vocal and choral music, some popular music, librettos, textbooks, instructional method books, and music periodicals. However, Braille music is produced by a relatively small number of institutions throughout the world.

A Braille model that is built on notions of BrailleCells and sequences of BrailleCells is used to codify musical aspects. The protocols of these coding rules are *not* standardized and there are still a few serious problems.

BrailleMusic is a sequential protocol with a defined grammar, much like serial communication through a RS232 serial interface. Starting from a graphical score, this means that a great amount of parallel data has to be 'packed' into sequential packages with a consistent structure. This can be achieved when using the WedelMusic Accessible Music Module to produce BrailleMusic (see Section 16.2.1.1.1).

16.2.1.1.1 WedelMusic

The WedelMusic project (http://www.wedelmusic.org), which ran from 2001 to 2003, provided a fully integrated solution that aimed at allowing publishers and consumers to manage interactive music: that is, music which can be manipulated (arranged, transposed, modified, reformatted, printed, *etc.*). This is done within a secure environment. The key components are a unified XML-based format for modelling music; reliable mechanisms for protecting music in symbolic, image and audio formats; and a full set of tools for building, converting, storing and distributing music on the Internet.

The interface for visually impaired people enabled blind musicians to have access to and share the facilities of the WedelMusic system in the same manner as sighted people. This implementation used many state of the art as well as conventional technologies like speech synthesis from text (widely known as TTS), Braille printing and technologies specifically developed by the WedelMusic project. The key element of this interface for blind people was the speech engine; this was based on Microsoft SAPI4 components (http://www.microsoft.com). A screen-reader program is used to provide a surrogate for the lack of visual feedback by conveying audible feedback to the user. For visual feedback to be fully and efficiently replaced, any activity is reported to the user. A Braille printing module was also developed to ensure that Braille music output is easily available, and in the format specified by the end-user. The importance of this development was considerable, as it allows far greater access to reading and creating music scores for visually impaired people.

An example of a first measure of a piece of music produced using the WedelMusic Accessible Music module is shown in Figure 16.1.

The BrailleMusic transcription can be described using Figure 16.2.

An example of the second measure after a navigation action by the end user to move to the next measure is illustrated in Figure 16.3.

Since the measures are represented as separate though associated objects, one choice is to interpret only these measures. In other words, the perceived context can be limited to only the current measure. Invoking the Braille interpreter on higher-level musical collections permits BrailleMusic interpretations of higher-level contexts. A simple example of such usage would be interpreting a page of BrailleMusic.

Figure 16.1. Navigation action through a musical score with additional real-time output to BrailleMusic

Figure 16.2. Measure 1, BrailleMusic explained

Figure 16.3. BrailleMusic interpretation of the second measure. Please note that the key information is missing, because the next measures contain no key information. In this particular set of interpretation settings, only key changes are presented

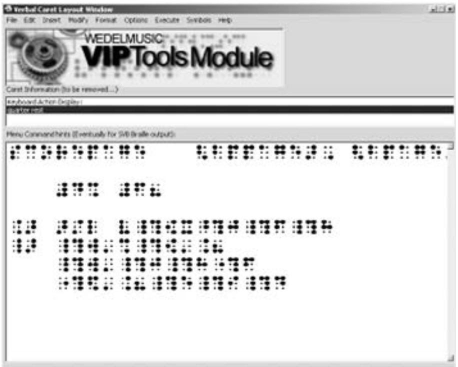


Figure 16.4. Real-time generation of a higher-level context description of a piece of music (in this case a page of measure)

An example of a page containing measures of a piece of music produced after selecting the appropriate menu item is shown in Figures 16.4 and 16.5.

The lower part preview of the BrailleMusic generated in this particular case is explained in Figure 16.6.

Figure 16.5. Measure 2, BrailleMusic explained

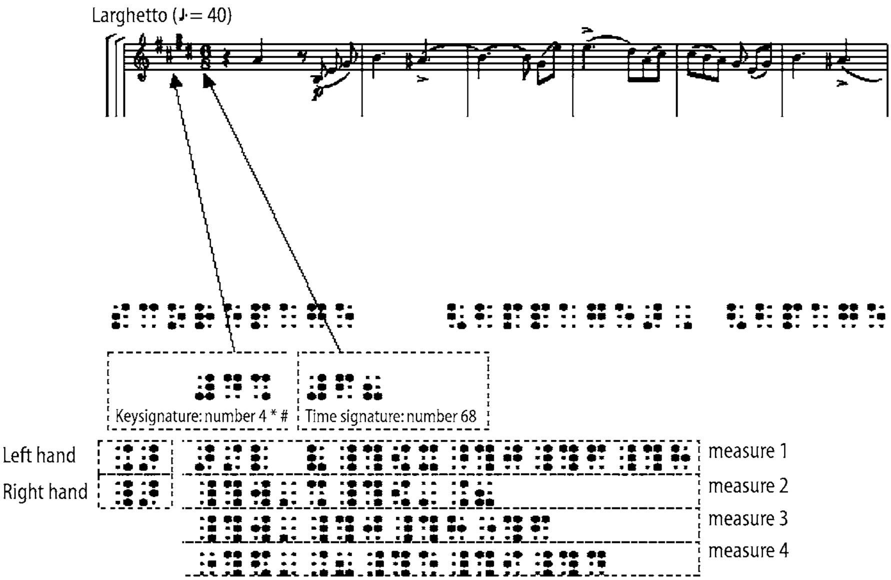


Figure 16.6. Higher-level musical information represented

16.2.1.2 Talking Music

This section firstly gives a general explanation of Talking Music followed by an explanation of an implementation of Talking Music using DAISY.

Reading Braille music is a complex task and, for those people who become impaired beyond the age of 20, it is extremely difficult to learn Braille. There is then a problem of finding a suitable teacher for Braille music, which can be difficult in some countries. Braille music can be seen as a specific dialect of Braille, so in theory anyone who can read Braille can learn to read Braille music. For those who *can* learn Braille music, there are often significant delays in receiving music scores, as the highly specialized production process is both time-consuming and expensive. Indeed, in many European countries the service either does not exist, or has been discontinued for reasons of cost or a drop in demand caused by fewer people learning Braille music.

One of the most popular alternatives to Braille music is Talking Music. Talking Music is presented as a played fragment of music, a few bars long, with a spoken description of the example. This spoken description describes the musical information using language which is familiar to musicians (*i.e.* fourth octave E quarter note *etc.*). The musical examples are now produced automatically by software and as far as possible with a non-interpreted version of the music. All notes are played at the same volume, at the same velocity, and all notes of a specific duration have exactly the same duration. The musical example is there to provide the user with a context for the Talking material, making it easier to understand and play. The musical example is non-interpreted, meaning that there is as little expression and interpretation in the musical meaning as possible. This affords the Talking Music user the same level of subjectivity to the musical content as the sighted user.

The most important part of the information is the name of the note. All the other information should be centred around this. Some of the additional musical information should be placed before the name of the note: some of it should be given afterwards. The guiding principles are the way in which the musical information will be processed, the duration of the musical information, and the importance of the information. For instance, there must be an indication which octave the note belongs to before the name of the note is said so that the listener can prepare some motor movements before playing the note. When, as an illustration of the second principle, a dynamic sign and a bow are mentioned at the same position in the score, the one with the longest duration should be mentioned first. Third, not all the information is equally important for everyone. For example, fingering is additional information, and is mentioned after the note and its duration.

Naturally, it is preferable not to describe every facet of every note, as some data are implicit. Having established a basic set of rules within DEDICON (formerly FNB, Dutch Federation of Libraries for the Blind) with several users, it was then necessary to consider ways of limiting the volume of information without losing any critical data. This led to the use of abbreviations. For example, the octave information does not need to be included for every note. If the interval between two consecutive notes is a third or less, then the octave information can be excluded.

In the same way, if the following note is a sixth or more above the current note, octave information has to be included. If the following note is a fourth or fifth above the current note, and the notes belong to different octaves, octave information does have to be included. The abbreviations adopted were the same as those employed in Braille music.

16.2.1.2.1 The DAISY Standard

Talking Music scores make use of the DAISY standard for Talking Books. The DAISY consortium (http://www.daisy.org) developed the DAISY format DAISY 3.0/NISO z39.86 (http://www.daisy.org/z3986/default.asp), which allows data to have an underlying structure for navigation reasons. The main application for this technology is in Talking Books for print impaired people. There are now hardware players for navigating DAISY structures and these are often supplied free of charge to visually impaired users in many countries.

DAISY is a key international standard to incorporate in accessible formats. Not only is it a structure with which impaired users are familiar, but it is also sufficiently well-established to be accepted in recognised hardware. A DAISY book contains many files, which are synchronised to provide the information and a means to navigate the information. A simplification of the DAISY format is shown in Figure 16.7.

There are several consumption means for the DAISY format. Hardware players are available which have simple interfaces that allow users to navigate from sentence to sentence, chapter to chapter or even book to book. Most visually impaired users in Europe have access to these. Other users prefer to use software players. There are various companies who design DAISY playback software.

A DAISY talking book is made up of fragments, which split the book into a structure. Each fragment contains an .mp3 file and a .smil file. The .mp3 file provides the audio information, while the .smil file provides the structural and

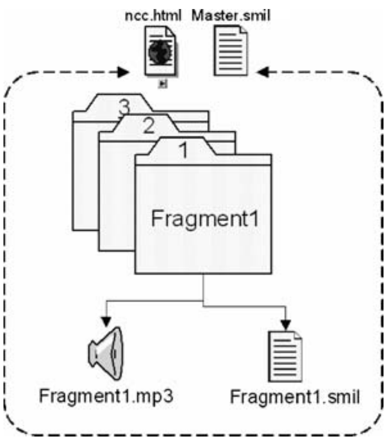


Figure 16.7. Basic structure of a DAISY Ebook

metadata information for that particular fragment. This essentially describes the content for each fragment. Each fragment has this information, and the .smil files can communicate with each other to provide different levels of abstraction. The collections of files are controlled and synchronised by a master .smil file and an ncc.html file, which provides the hierarchy for the structure, and is the file which is read by the DAISY player or software.

16.2.1.2.2 Talking Music Output

One of the guiding principles of universal design and Design for All guidelines (www.design-for-all.org) is that the print disabled should have access to the same information as the sighted, and that only the format in which the information is presented should change. For Talking Music, this means that everything on the page of a music score must be represented in a Talking format. Furthermore, this Talking format must be applicable to all types of music and instruments. By way of illustration, Figure 16.8 contains an excerpt form a Bach minuet.

The resulting Talking Music for measure 1 of the simple example above is shown below:

> The title is: Minuet in G. Composer: Bach.
> This piece contains 8 bars and consists of 2 sections. The piece is read in eighth notes, except where otherwise indicated.
> Key Signature: 1 sharp. Time Signature: 3 fourths time.
> Section 1.
> Bar 1.
> second octave g half.
> a quarter.
> in agreement with
> fourth octave d quarter.
> third octave g.
> a.
> b.
> c.

The description above shows the information within the Talking Music book taken from a DAISY Software player. The description starts with a header which

Figure 16.8. Bach excerpt – Minuet in G – BMV 114

delivers the document data about the score (composer, title, production notes), a description of the score (in terms of length and content), and then the key and time signatures. The description then continues to Section 1, where it describes the first measure note by note. The lowest voice is described first.

This information exists as audio files, and the text description is used for production use and explanatory reasons. The audio files are structured in such a way that the user can easily navigate from one section of the score to another, in order to build up an idea of what is in the score. The tree structure of a more complicated score taken from a DAISY Software player is shown in Figure 16.9.

A Talking Music book exists as a collection of files controlled by an .ncc file but to describe the structure it helps to think of a folder hierarchy. The Talking Music Book contains at the first level folders for the header information and folders for each section. The folders for the header information can be navigated through vertically in order to skip to the desired piece of information. There is then a midi representation of the entire score. The midi representation provides a computer generated playback of the music. This is usually played at the standard tempo, but the tempo can be altered in the production process if desired. The sections then contain a folder with the midi for that section (usually slower, depending on the complexity of the piece). The midi is then followed by a folder for each measure which in turn contains a folder for each voice. After the first level the score can be navigated both vertically and horizontally. This means that the user can jump from section to section, measure to measure or voice to voice. This affords the user the luxury of relating the information with its context in a way similar to a sighted user, who can overview the score at several levels of abstraction.

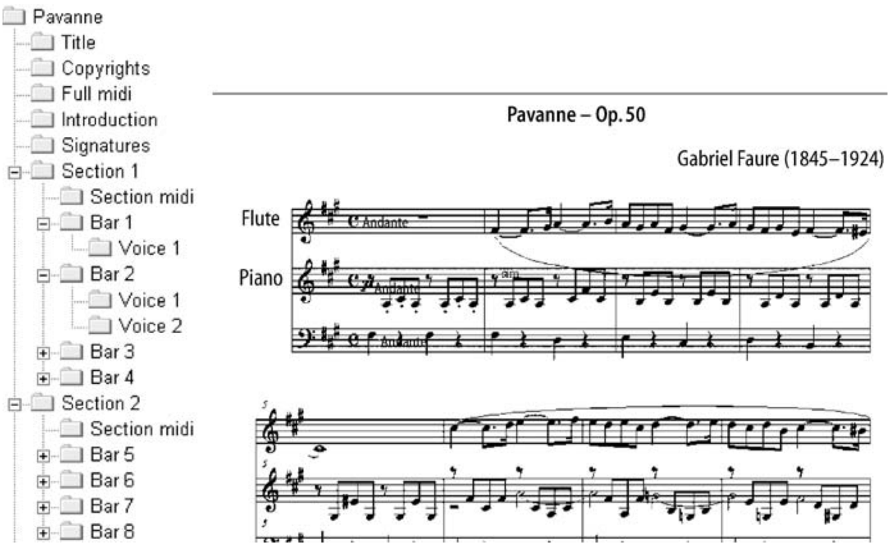


Figure 16.9. Talking Music Score structure and an excerpt from the score

16.2.1.3 Large (Modified) Print Music

A large number of the users of accessible formats are not blind. One popular accessible format for many years has been that of Large Print Music (or modified stave notation). Because this is based on traditional common western music notation (CWMN) it is very widely used by people with diminished sight who could once use CWMN.

The concept is simple; the music notation is made larger than normal, where the scale factor is relevant to the quality of the user's eyesight. The common way to do this is to take the A4 score, and to increase it to A3 using a photocopier. With the increasingly common use of computers as a composer's or typesetter's tool, the score becomes more adaptable to accessible solutions. Considering the previous example of the Bach minuet, the size of the zoom in a music notation package can be increased, and an example is given in Figure 16.10.

This is sufficient in many cases, but this resolution is the highest achievable before the pixels become visible, and the image becomes less useful. When using music notation editing software it is possible to optimise the score for large print versions. The notes can be typeset in such a way that the durations are placed evenly across the measure. There should be little or no overlapping information. As with all accessible solutions consistency should exist wherever possible; if dynamics appear below the measure they should continue to appear below the measure for the whole score. These guidelines make the score more viewable at small font sizes, and so should help the user in understanding the different elements of the data.

Large Print Music notation can benefit from the information digitisation process. It is important from an accessibility point of view that the traditional score is not modified, but that the information is presented in different ways. Interesting solutions would be possible if modern extensible methods were used and incorporated at the graphic rendering stage of the music notation production. One possibility in this field would be the use of scalable vector graphics (SVG – http://www.w3.org/TR/SVG/), where an XML-based description (http://www.w3.org/XML/) of the graphic is used and, because it is vector-based, the representation can in theory be used to increase the size of the rendering to any level. This extensible modularisation of the information also allows the different elements in the score to be represented in different ways. This is especially important for dyslexic users, as the variance in problems with information is greater, and personalised accessible formats may be a viable option.

Figure 16.10. Example of large print music notation

16.2.2 Technical Aspects

The system is a good example of accessible information processing. Through the use of standards, and an awareness of user requirements, an accessible music processing system has been created which can meet the needs of end-users. Such a system has more likelihood of being integrated into existing workflows and supply chains, as it has recognisable connection points to frameworks which are already in place.

In order to achieve this robust accessible design, there are few technical aspects which should be taken into account. These are tackled below. It should not be forgotten that one of the most important aspects of accessible design is meeting the (often specialized) user needs. This can only be achieved through user testing and consultation throughout the design process.

16.2.2.1 Hardware Aspects

These will serve as 'hardware parameters' (see Table 16.2). The aim is to abstract these aspects into proxy entities. This enables OS dependent hardware devices and their associated services and drivers to be mapped to addressable entities in the accessible music production layers of the system and also the end-user consumption layers.

16.2.2.2 Software Aspects

The aim is to provide abstract proxy services that render the operating system dependent features addressable. Table 16.3 shows software components for accessible systems.

16.2.2.3 Usability Aspects

The aim is to collect usability and user requirements (see Table 16.4) into proxy services that grant access to these features and requirements and serve them as parameters. A user model will emerge. This user model should be an *extensible* user model. Every user is an individual and, moreover, requirements and preferences might change. This dynamic behaviour should be available in all proxy services.

Table 16.2. Hardware components which can be used in accessible systems

Computer Input Devices	Computer Output Devices	Hardware players
Mouse	Monitor	CD player
Keyboard	Speaker/headphones	DVD player
Data input (drives/memory/network)	Braille bar	DAISY player
Joystick	Haptic devices	Smartphone/PDA
Gestural (camera)	Braille embosser	

Table 16.3. Software components that can be used in accessible systems

OS specific	Visual specific	Audio specific	Application specific
Braille bar driver	Graphical desktop environment	Text-to-speech	Music production
Speech engine	Shell	Midi	Music playback
	Screen reader	MPEG/wave device	Music archiving/library
		GUI to speech	CD/DVD production
			Batch production

Table 16.4. User aspects which can be used in accessible systems

Navigation	User aspects	User preferences	Production limitations
Context/location information	Level of impairment	Level of personalisation	Time
Action feedback	Level of environment influence	Overlap with other solutions	Volume
Persistency – session management	Level of software experience		Skill availability
Consistency	Level of hardware experience		Cost
	Age		Batch size
	Comprehension		
	Language		

16.3 Some Recent Initiatives and Projects

This section gives an overview of some of the recent initiatives and projects in the field of accessible music software for print impaired people.

16.3.1 Interactive Music Network

The Interactive Music Network (http://www.interactivemusicnetwork.org) is a centre of excellence created to assist the music industry to enter the marketplace by drawing on the assets and mutual interest of the different actors in using technologies, tools, products, formats and models. This thematic network, funded by the European Commission Fifth Framework IST Research Programme (http://www.cordis.lu/fp5/home.html), addresses chiefly the area of music coding, as there is still no accepted standard for music encoding. The main aim of the network is to bring music into the interactive media era by setting up a series of activities such as:

- Creating a web gateway.

- Publishing specific periodic reports and surveys.
- Producing integrated solutions and guidelines proposing standards.
- Producing guidelines for the adoption of current active standards.

The main idea of the Music Network was to create a community to bring the European music industries and content providers together with research organizations and emerging technologies. The Music Network set up seven different working groups on the most important aspects, in order to bring music industries into the interactive multimedia music age. The working groups integrated their results in order to produce new standards or guidelines to the existing standards in the area of coding multimedia interactive music according to the needs of the music industry. Project partners, multimedia companies and research institutes have produced trial applications customising their present applications in order to demonstrate the effectiveness of the provided network support.

The Music Network Working Group on Music Notation was set up to come to an agreement on one definitive music notation format. Music notation is not only related to music printed on a sheet, as notation will also be accessed using different kinds of devices, from a PC to a UMTS terminal, from the classical printed music sheets to the electronic lectern. Music notation is related to other media, *e.g.* audio/videos of performances, documents related to the piece of music or to the composer (biography, critical notes), images of the author, performance, director, *etc.*

Another important aspect is the level of interactivity that can be achieved only by using a symbolic description of music notation. Images of the music score (TIFF, GIF, PDF files) cannot be transposed, modified, or annotated. However, interactivity presents copyright problems, and so digital rights management (DRM) support is needed to give the publisher control over the use of the music.

At present, no open music notation representation format has been widely adopted as a standard. And no format supports multimedia aspects and interactivity under DRM. A music notation format allowing multimedia integration and interactivity under a digital rights management framework is needed. Work has been started by the MPEG Ad Hoc Group on Symbolic Music Representation (http://www.interactivemusicnetwork.org/mpeg-ahg) on this subject.

The Accessibility Working Group within the Interactive Music Network brought together various stakeholders in accessible music during the course of the project. Several interesting initiatives were presented during technical workshops, and panel sessions were held to discuss some of the requirements involved in accessible music, education, production and technology. These involved end-users who were invited guests in the sessions. The tangible outcomes from this working group can be seen in the continuing work on accessible music notations within MPEG.

16.3.2 Play 2

The PLAY 2 project (http://www.dodiesis.com) was an EU-funded project which aimed at facilitating access to digitised music for visually impaired people. Perhaps the most innovative aspect was the Braille music editor. The Braille music editor (BME) is conceptually a real file editor, with normal facilities like input facilities, reviewing facilities, output, saving, cut and paste, print *etc*. The difference is that BME deals with Braille music notation. The input code is the Braille music notation, as coded in the "New international manual of Braille music notation" (Krolick 1996). In other words, the blind user types in the score as on a Perkins writer, following the rules of Braille music notation. Reviewing is done using the four arrows, and Ctrl+arrows, much like normal reviewing in text environment. Acoustic feedback happens simultaneously in three different ways: Braille (on Braille display), midi, and through a vocalised musical element. The end-user may print the score either in Braille, or in ink print.

The BME can accept data not only from a keyboard, but also from a scanner scanning Braille paper. The BME can also accept data from an ASCII file, which corresponds to a Braille printable file. This feature turns out to be particularly useful for proofreading. With this back translator the sighted transcriber is able to view directly and personally the result of their work and this reduces time and costs in Braille music production.

16.3.3 Dancing Dots

Dancing Dots Braille Music Technology, Limited Partnership, (http://www.dancingdots.com) was founded in 1992 to develop and adapt music technology for the blind. In 1997 Dancing Dots released its first product, the GOODFEEL® Braille Music Translator. Bill McCann, Dancing Dots' president and founder, sees GOODFEEL® as the first in a series of high-tech tools to harness the power of the personal computer for creative people with disabilities. Bill McCann himself is a blind musician and programmer.

The programme is fast and flexible. A single part, selected parts, or an entire score can be Brailled in seconds. It features integrated literary Braille translation for most western languages so both the words and the tune can be Brailled. GOODFEEL® supports the various formatting styles used around the world and can be configured to conform closely to the formatting conventions used for Braille music production in its country of use. The programme runs on the latest versions of Windows: Windows XP Home, XP Professional, 2000, NT, Millennium, 98 and 95.

With GOODFEEL® combined with a few mainstream products, any sighted musician can prepare a Braille score without needing to be a Braille music specialist. Blind users can make sound recordings and print and Braille editions of their compositions and arrangements. Music scanning software can be used to speed data entry. Dancing Dots offers a variety of products including integrated software training modules, and assistive technology such as JAWS (http://www.sightandsound.co.uk/pages/JAWS.htm) and the Duxbury Braille Translator (http://www.duxburysystems.com/).

16.3.4 Toccata

Optek Systems (http://www.opustec.com) support and distribute a wide range of electronic products that assist people with a disability throughout Australia and New Zealand. One of their products is the Braille Music Software Toccata (http://www.opustec.com/products/toccata), a tool for parents, teachers and Braille transcribers to produce quickly accurate Braille music in a variety of formats either from scanned sheet music or using Toccata's Notation Editor.

Toccata is a very simple to use Windows program giving various options that greatly reduces transcription time. Using the mouse or keyboard, notes can be placed into Toccata's Notation Editor to create music of any complexity. As notes are placed, they are heard through the computer speakers at the same pitch and duration as the note selected. The music can be played back in real-time. A separate window shows the translated and formatted Braille, which can be directly sent to a suitable embosser.

The Braille can be instantly reformatted in common formats, *e.g.* Bar-over-Bar, Section-by-Section, and a wide range of optional styles can be incorporated, *e.g.* Show Hand Signs, Lyrics in Grade II contracted Braille, and so on. It is also possible to edit the Braille directly with six key inputs, and at all times the Braille and Music Notation Windows are synchronised to assist with proofreading. Alternatively, an entire music score can be loaded as a MIDI (http://www.midi.org/about-midi/specshome.shtml) or .NIFF file (http://www.musique.umontreal.ca/personnel/Belkin/NIFF.doc.html) or scanned into a computer *via* a flatbed scanner, displayed on the screen and then edited using the mouse.

16.4 Problems to Be Overcome

In the process of entering, transforming and getting the modified music out of the computer in any format using any device, three processing stages can be distinguished. The *input stage* collects all activities that are aimed at interfacing the real physical world to the virtual computer world. Inside the computer world all kinds of processes take place. An important notion about this virtual world is the fact that theoretically *no* loss of information may occur. As long as information and all related entities remain digital (*i.e.* virtual) there is an opportunity to model the flow of information and the processes that transform and govern the information in a non-destructive way. This is instead of models that take only the information that is used for the primary application of the content, *i.e.* information is removed as it is deemed additional. This stage is called the *representation stage*. The last *output stage* is the phase in which the representation stage is transformed in order to re-enter the physical world. Re-entering the physical world means adapting the represented information and the representation of the structure in which the information is stored to meaningful and comprehensible bits of information. These bits of information are communicated through hardware devices. Hardware devices that enable end-users to touch, hear, see or take in information through any combination of these sensory stimuli.

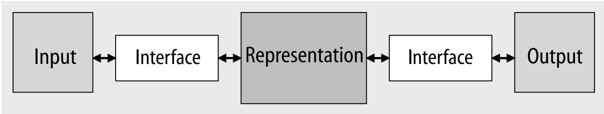


Figure 16.11. Three-layer information processing

The end-user makes the final interpretation of the bits of information that are offered through these channels. Such a three-layer model can be seen in Figure 16.11.

Using the separation into the three processing stages outlined in Section 16.4, that is input, representation and output, this provides us with an abstract framework for the identification and structuring of the various components that interact during the process of perceiving, representing and interpreting information. An important feature of such a separation is that it provides a level of *unification* in approaches of tackling information provision to *any* end-user. This will enable—in due time—equal treatment of common users and special needs users, such as visually impaired users, dyslexic users, elderly users, *etc*. There are also reciprocal opportunities. Due to the extreme requirements of the special user group projects onto information processing systems, the quality of the common requirements will improve. The 'resolution' of these requirements will be improved. Moreover, because both types of requirements may be addressed and fulfilled, any type of end-user will benefit greatly from the improved performance and quality of the information processing and delivery framework. Given this unified approach, a sustainable and generic solution might be found.

16.4.1 A Content Processing Layer

Initiatives are needed to introduce accessibility into the general market. End-users, and producers as well, who work in the accessibility arena are frequently confronted with the 'so what' attitude. Accessibility is regarded as a niche market, and a niche market that does not apply to *healthy* people. This is a very short-sighted vision, especially if one bears in mind the fact that everyone grows old, might have vision that deteriorates, might have hearing that deteriorates, might have a speed of understanding that might not be sufficient to keep up with the exponentially growing possibilities of technology and so forth.

What is needed is the means to explore various combinations of processing configurations. Using this possibility of exploration, the process can be optimised and while doing so production results can be retrieved. 'Managed' exploration also permits specialization of the information processing configurations to alternative uses for specialist needs *in parallel* with each other. Both specialist instantiations of the processing configurations will have a shared abstract framework that describes them in relation to each other: the explorative framework, or the MPEG21 framework (http://www.mpeg.org/). An example of such a framework is shown in Figure 16.12.

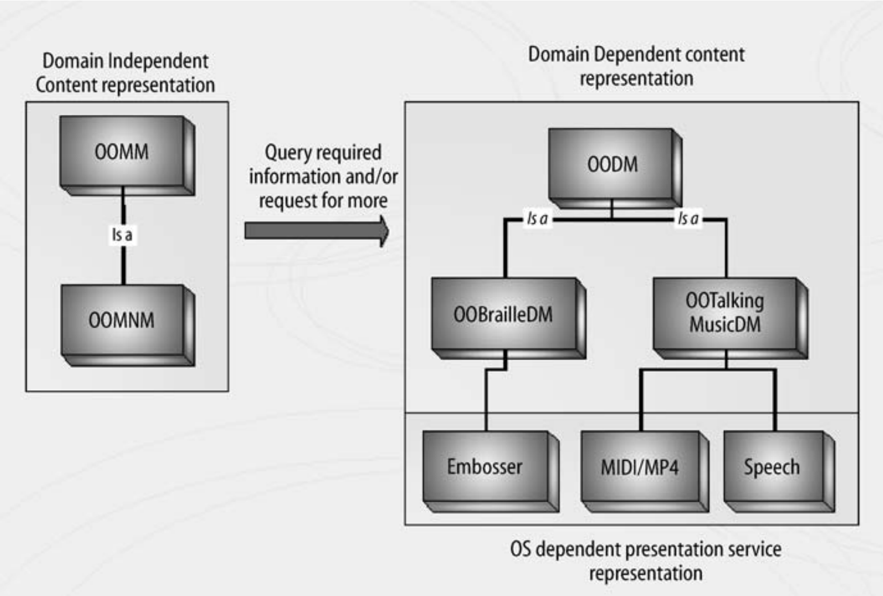


Figure 16.12. Accessible music information processing framework

Figure 16.12 handles the representation block of the three-block framework. For convenience, it is split into domain independent and domain dependent content representation layers. These can be seen as separating content from presentation.

One object oriented music model (OOMM) is created. There is also a specialization of this model called the object oriented music notation model (OOMNM). The domain dependent document modelling elements of the system then attach to these elements using queries. The information retrieved is then used to form the presentation of the content.

The generic presentation model is the object oriented document model (OODM). In the case of the system in question this is split into an object oriented Braille document model and an object oriented talking music document model. These presentation objects then in turn feed out into the relevant output modules for the supported output formats.

16.4.2 Standardization of Accessible Music Technology

For accessible music to reach a broader audience, and with that simplify the process of retrieving accessible music for those who require or prefer it, there exists the need to integrate notions of accessibility into widely used formats and methodologies. If the fundamental issues of accessibility become part of an *agreed standard* that is applied to various media domains and process description domains, the knowledge *about* accessibility is integrated into the standardization documentation. Because of this integration into the standardization process and consequently the implementation of it in the form of applications that conform to this accessible standard, a level of education is achieved.

Accessible applications ‘teach’ end-users what accessible information is. Standardization process documentation in the form of system design notions and APIs ‘teach’ developers, architects and programmers what accessible information is. Implementation schemes and market targeting mechanisms ‘teach’ project managers and entrepreneurs what accessible information is. This continuous process of *hints towards accessibility* will stimulate any kind of user to think about accessibility—be it consciously or otherwise. The end result might be a big leap towards sustainable and accessible communication based on *use* rather than simply specification.

16.5 Unifying Accessible Design, Technology and Musical Content

This section builds on the key issues identified above and illustrates techniques which have been used for applying this integrative approach to Braille Music and to Talking Music.

16.5.1 Braille Music

16.5.1.1 Creating a Braille Decoder Module

All BrailleMusic related classes and components can be collected in a BrailleMusic package. All individual Braille modules and classes are grouped into six subpackages, as in Table 16.5.

The exact structure of the various packages can change and more sub-(sub-)packages can be defined in the future. The Editor package is a dummy package that is used to illustrate the relations between the BrailleMusic packages and object oriented music model (see “Producing Accessible Multimedia Music”. Crombie *et al.* 2003c) and editor-level or application-level packages.

Figure 16.13 shows all currently defined packages and their relations. These packages will be described in greater detail afterwards.

Table 16.5. List of Braille modules

1. Braille subpackage
2. BrailleDocument subpackage
3. BraillePrinter subpackage
4. BrailleMusicInterpreter subpackage
5. BrailleMusicInterpreter. Rules subsubpackage
6. Editor subpackage

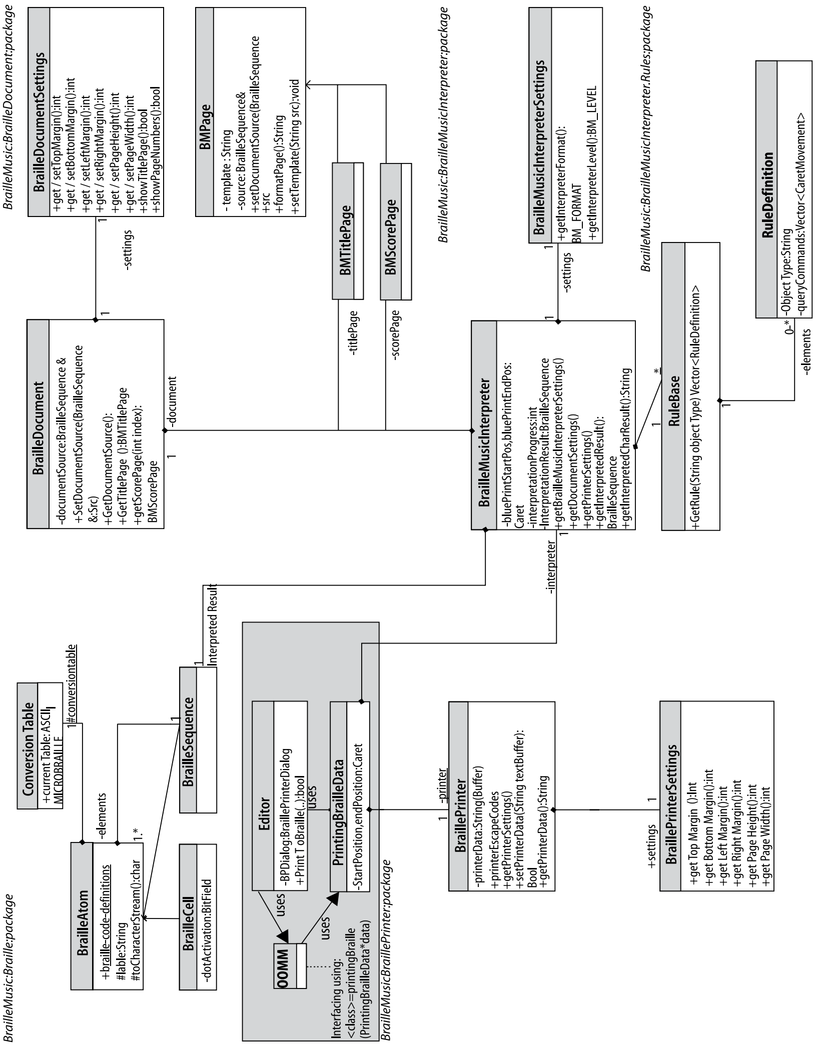


Figure 16.13. Braille decoder system architecture and components

One of the goals is to create as much decoupling between different modules as possible. This is in order to achieve a clear description of the complex nature of the required procedures and processes. Another reason to strive for maximum decoupling between modules is the 'living' nature of the BrailleMusic notation standard. This means the language can be expanded or adapted at any time. Therefore, a dynamic and configurable BrailleMusic rules model is required. The modelling of Braille classes should be flexible in order to facilitate separate modelling of Braille 'content' and Braille formatting. By transforming the Braille model to ASCII characters, the use of generic Braille printers and re-use of the same Braille information on-screen is possible without sacrificing modelling flexibility. It is also important to be able to retain groupings of Braille cells in order to keep the sentence structure intact.

16.5.1.2 Braille Element Classes

In order to fulfil the goal defined above, a hierarchy of Braille related classes are introduced: BrailleAtom, BrailleCell and BrailleSequence. BrailleSequences are sequences made of other BrailleSequences and BrailleCells. BrailleCells and BrailleSequences share several constant definitions and certain procedures, which are collected in the BrailleAtom class. A BrailleAtom class is an abstract super class of BrailleCell and BrailleSequence. These relationships are shown in Figure 16.14.

The BrailleMusicInterpreter produces these Braille related classes. A BrailleMusicInterpreter class represents the notion of BrailleMusic Interpretation, meaning the coupling between musical structure and its context and BrailleMusic transcription. This coupling is achieved by means of Interpretation Rules that are related to specific musical object types and complexes of musical object types. The BrailleMusic Rules form specific queries directed to the object oriented music model in order to produce a Braille output. The exact definition of rules can be specified by a BrailleMusic interpretation expert (a composer or teacher of BrailleMusic).

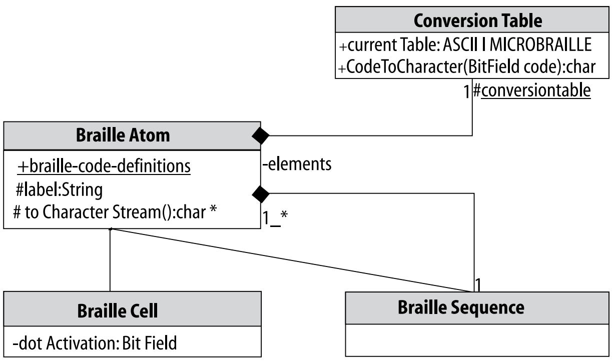


Figure 16.14. Braille class relationships

16.5.1.3 Braille Printing

In order to facilitate expansion of the number of BrailleMusic interpretation rules, the representation and the management of these rules are modelled with the Rule Base and Rule Definition classes. The Rule and Rule Definition classes model the required musical queries. That is, they specify which rule needs what musical information and which procedures should be executed to obtain that required musical information. The obtained musical information is used to create a BrailleCell/BrailleSequence hierarchy tagged with the relevant labels.

Non-musical parameters can also be of importance in the interpretation process. For instance, the number of columns a Braille printer can print influences the interpretation of a melody: if little space is available, the expert will try to achieve an interpretation that is as compact as possible. To enable the interpretation model to incorporate these non-musical parameters all relevant settings, such as Document Settings, Printer Settings and BrailleMusic Interpretation Settings, are modelled in specific classes. All these classes can provide non-musical information needed for interpreting BrailleMusic. These classes can also handle relations between the different Settings classes, such as, for example, BraillePrinter.TopMargin and BrailleDocument.TopMargin.

The Braille classes are capable of converting the hierarchy of BrailleCells and BrailleSequences into a stream of ASCII characters. The exact choice of ASCII characters that are associated with specific BrailleCell configurations (*i.e.* Braille dot activations) is dependent on the choice of the current Conversion Table. A conversion table for creating ASCII-table-based or MicroBraille-based character streams is available. These character streams can be directed to a BraillePrinter class for printing to Braille or can be directed to a screen resource. In this way the character stream can be read by screen reader software, such as JAWS.

16.5.1.4 BrailleMusic Interpretation Process Description

The BrailleMusic interpretation process can be described as a large set of rules. Rules that apply to the musical information that is available through the music model that is used in the music program. The music model provides primitive building blocks that allow the programmer to create grammatical structures of musical material. Once this grammatically correct description of musical material is available, additional translations—or transformations—can be applied to it. Transformation of the music model that contains a 'grammatically' structured description of the musical material in digital form to 'grammatical structures' that make sense from the perspective of Braille music can be described as a language interpretation process. The definition of this interpretation can be described using the same technology and techniques that were used to conceive the non-specialized music model we described before.

When describing the common music to Braille music interpretation steps using digital components that will embody this interpretation process, there is a need to group these grammatical rules into meaningful groups. The grouping of interpretation rules employed in the "New International Manual of Braille music notation"

(Krolick 1996) is found to be useful for this purpose. This manual is in widespread use, so we can safely assume this grouping scenario to be familiar to the accessible music production domain. Sticking to the structures conceived and used in the manual will hopefully make the transition from the printed version of the manual to the digitally embodied version of the manual more intuitive. Furthermore, the manual already provides an almost rule-based description of the interpretation steps required to transform common Western music notation to Braille music. The manual does this by means of providing lots of examples that explain the interpretation process step-by-step. The general structure of the axioms described in the manual is similar to: ***if <current-pitch> is preceded by <another pitch> describe <current-pitch> as xxx***. The only additional step is the description of the building blocks of this interpretation process in digital components that are able to execute the individual steps of the Braille music interpretation process.

The Braille music interpretation rules can be grouped in six interpretation levels, as shown in Table 16.6.

The decision was made to approach the interpretation process in a bottom-up manner. In musical modelling terms, this means that we start from the individual notes and travel up in the hierarchical level of the musical model after finishing the note level. This approach ensures consistency of individual note descriptions in Braille music and additionally it introduces the possibility of allowing the next level of processing (that is of a higher abstract level) to incorporate the already interpreted lower level note descriptions. This means that the more advanced musical processing that involves notions like slurs, braces *etc*. that in turn involve groups of notes can be described in separated packages of musical interpretation rules. So the individual rules related to each level will be implemented in ascending order. That is, start with level 1 rules and continue with level 2 rules. All rules related to level 1 are required for level 2. The same procedure applies for the other levels. The exact number of rules per level will be increased in the future. The goal is to start with a basic set of rules per level that enable us to interpret and print basic music sheets. Each iteration that is applied to check the consistency of musical interpretation can be used to fine-tune the appropriate set of rules, without inflicting too much damage to the other interpretation rule groups. In the end all of the groups are expected to be fine-tuned and more attention can be given to the *interplay* between the groups of interpretation rules.

Table 16.6. Interpretation levels for Braille music

1. Level 0: Cover level (title page)
2. Level 1: Pitch level
3. Level 2: Chord level
4. Level 3: Phrase level
5. Level 4: Lyrics level
6. Level 5: Repeat/segment level

In the approach taken the RuleBase class will function as a central registry that administrates all available Braille music interpretation rules based on the notions that were described above. The administration system is based on the abstraction level and assigned musical information class type of the individual rules. The administration module that governs the Braille music interpretation process provides efficient ways of querying the registry of music interpretation rules that includes embedded knowledge about the structure of the musical objects that are used to describe both common Western music notation and the Braille music rendition of this. The entries contained in this registry of musical interpretation rules, can be modified dynamically. If the definition of the rule needs to be changed, it can be changed. If an additional rule needs to be added to be able to cope with more complex kinds of musical material, more rules that focus on these newly required features can be added to the registry.

The exact definition of a rule typically is represented by a consistent set of computer code that embodies the individual steps required to be taken to gather the required musical information. Normally such sets of computer code are hard-coded into the system. If in such a form changes were made to the definition of the interpretation grammar, the whole computer application would need to be recompiled. This is a very labour-intensive way of incorporating flexibility into a language interpretation system. The approach taken in the accessible music production tool described includes strategies to relieve the need for hard-coded computer code, by means of a scripting language. The computer code that manifests the interpretation process operates on a meta- level, meaning that it 'knows' what a grammar means and how it should apply certain sets of interpretation directions to certain kinds of musical constructs. The actual content of such a step is taken from a persistent XML file, with accompanying XSLT file, that describes the exact steps and its associated queries required to complete successfully the individual Common Western Music Notation (CWMN) to Braille Music Notation (BMN) transformation steps. Ultimately this allows dynamic changes to be made to the definition of the grammar used to interpret CWMN to BMN, whenever and by whoever required.

The latter feature not only provided a useful tool for fine-tuning the common western music notation to Braille music interpretation process. It also allows for multiple interpretation descriptions to be present in the RuleBase. For example, it is possible to instruct the registry to contain a set of groups of interpretation rules that are most efficient for Braille music production in the Netherlands, alongside group interpretation rules that are most efficient for Braille music production in France.

16.5.1.5 BrailleMusic Page Formatting

For the formatting stage, some assumptions are made in relation to the interpretation process. The interpretation process is responsible for the deduction of logical BrailleMusic Sections, be it intelligently deduced or low-level chunked.

The parameters for document formatting are shown in Figure 16.15. The page formatter will request BrailleMusic sections and expects to find BrailleSequence

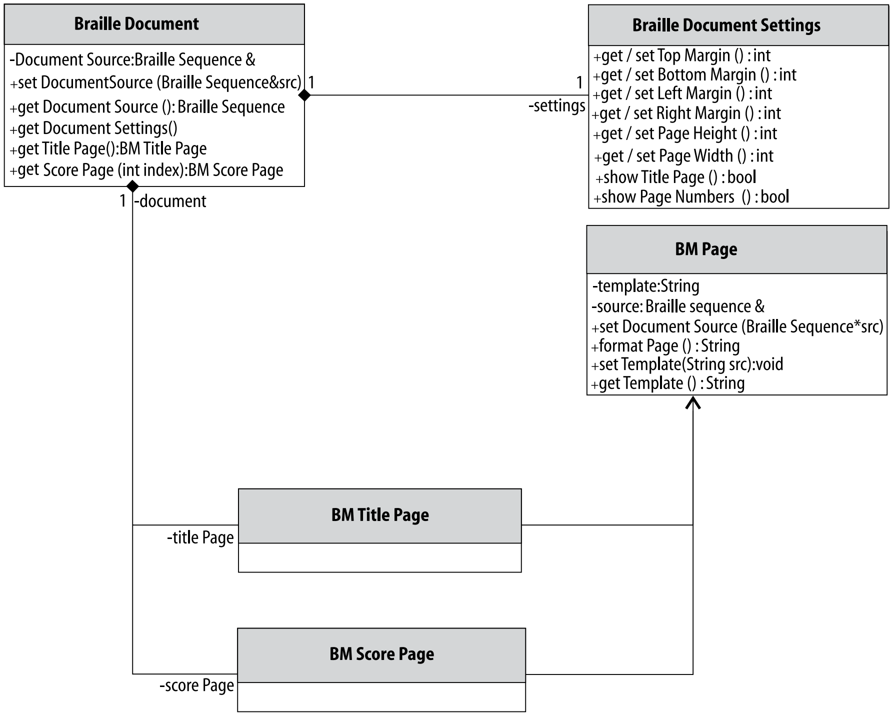


Figure 16.15. Parameters for document formatting

objects with labels relating to sections. The page formatting process can be described in the following steps:

1. Retrieve the preferred document layout properties, such as: show page numbers, show title page, and other available properties. A default page template is already initialised but this can be overridden, thus making it possible to format the page in another layout.
2. If required, invoke the TitlePage object's format method. This TitlePage object will in its turn invoke all required methods and resources to construct a title page containing title, author, publishing date and other information. The title page contains a predefined template. This template captures the page's layout and provides 'slots' or 'keywords' that will be used to insert labelled information into the required location in the layout. The result of this page's formatting (that is applying the page's template) will be stored in its character buffer and will be kept there until needed elsewhere.
3. Depending on the amount of Braille information available (or in other words: the length of the score) an iteration loop is started to either create ScorePage objects containing section information or map over the Braille result using a ScorePage object as a dynamic template. The formatting procedure for a ScorePage is the same as the one for a TitlePage. The difference between

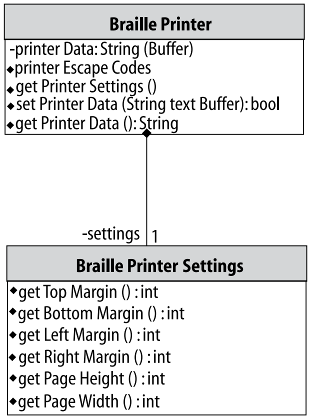


Figure 16.16. Classes that control the Braille printing

the two templates is the use of different insertion tags. A table of valid insertion tags will be available soon. Since the templates request named BrailleSequence objects that contain the syntax of the Braille chunk, line breaking and other separation issues can be handled here. If the character output of a BrailleSequence would not fit on the same line, a line feed can be inserted to force the character output to be on the same line. If a ScorePage object is successfully formatted, it will keep the result in its own character buffer until requested later. The latter procedure depends on the exact implementation of the ScorePage generation iteration loop.

After the steps described above, all 'raw' character information needed for printing to either an embosser or printer is available. All information is tailored to the preferences defined in the various objects' settings. The final step in the Braille printing procedure will transmit the required/preferred parts to the spooler.

The final step in the Braille printing procedure invoked from the PrintBraille method associated to the 'Print to Braille' menu item is the transmission of the interpreted and formatted character output to the printer spooler. The printer spooler will handle the actual transmission of the data, as shown in Figure 16.16.

In the last execution step in the PrintBraille method the character buffers of the BrailleDocument's TitlePage and various ScorePages are copied into the BraillePrinter buffer (or streamed to the BraillePrinter buffer). The BraillePrinter object will issue all required commands to send the buffers to the spooler. Notifications (for instance: error messages, print job finished, *etc.*) are handled by means of the BraillePrinter object.

16.5.2 Talking Music

Talking Music was originally produced by taking the music score and writing a 'script' for a trained reader which described the musical information within the score. The results proved very popular with end-users, but the process had several underlying problems. The transcribers of the music who wrote the script used slightly different protocols, as many of the aspects were open to subjectivity. This resulted in users receiving inconsistent output, while it has been outlined above that consistency is a core requirement of Design for All.

From a logistics point of view, the production was very time consuming, with a transcriber, a reader and then a technician to parse the chunks of audio into a suitable structure. A further loss of productivity occurred because the music, being essentially a unilingual representation, was fixed to one language. The process cannot take place successfully unless it is automated.

One of the major strengths of an automated method of development is that the output can be produced in almost any format. The creation process is dynamic so the user has almost unlimited control of how and what the program produces. This increases the use of the technology and meets the needs and preferences of a larger body of consumers. In particular the preferences regarding musical detail are very important. In order to meet these needs effectively a way has to be found to model them. A first step in modelling these needs is to meet the requirement for a flexible means of music representation. This flexible music representation method should at least encompass all the musical input information that is provided at this moment. It should also be flexible enough to allow compatibility with any input or output format or user need added in the future. This resulted in the specification and development of an object oriented music model and all related components.

If the representation framework for parsing the musical information described above for the Braille music requirements (Section 16.5.1) is robust enough, then in theory any output format can be constructed from the object-oriented model. Using the same model facilitates much more personalisation for such requirements as automated Talking Music creation.

As outlined above, Braille music has several varying requirements from country to country as to how it is used. Using the same methods used to solve this problem, a Talking Music system can be created which allows not only non-destructive modifications to take place on the score but extra languages can be easily created, so it becomes simple for any producer to output in several different languages. These requirements all come under personalising user requirements; the scalability required to create multiple language outputs has the secondary advantage of allowing users to specify their own grammar in the Talking Music Protocols used.

In order to achieve the level of interoperability required for Talking Music, and ultimately to allow the user to navigate the score, a clearly defined output structure is required which each score can conform to. A simplification is shown in Figure 16.17.

The document contains two main sections, one with the Header information (Metadata) and one with the literal music information. The elements that make up these various repositories of information allow the output requirements to pick and choose what complexity is required.

A closer look at the parsing of the musical information is shown in Figure 16.18.

The musical information becomes hierarchical in this simple example, but it is important to note that the entities involved in this framework are represented in such a way that music can be presented in a non-hierarchical way or in a dif-

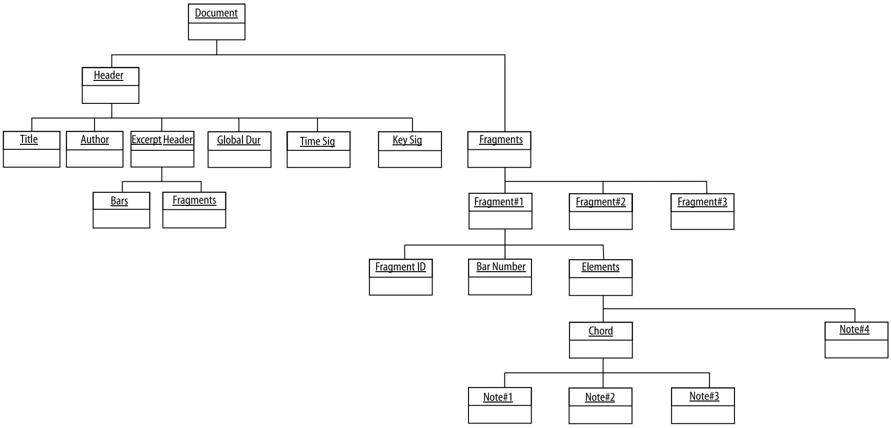


Figure 16.17. Simplified graphical representation of a Talking Music structure

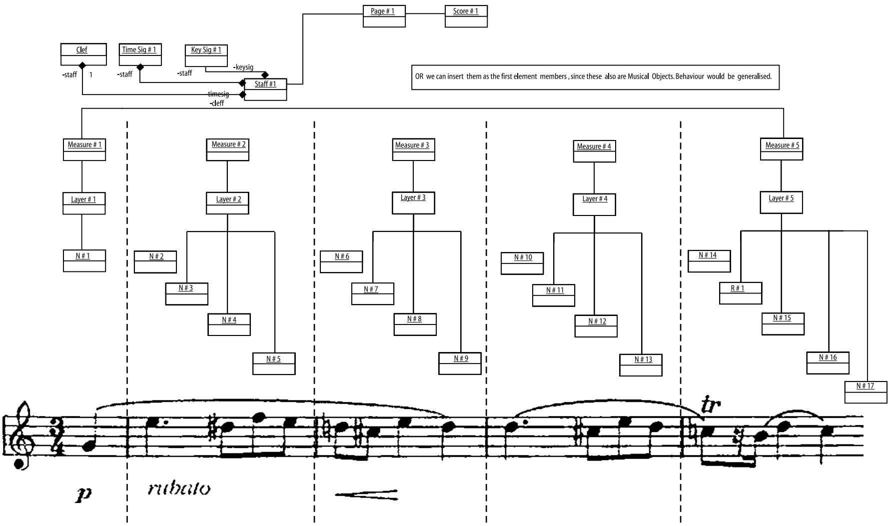


Figure 16.18. The parsing of music to the note level in the Talking Music framework

ferent hierarchy. This multifaceted output can be made possible only through the transcription of 'intelligent' content. If the music is created in such a way that all options are left open and no information is left out, then transcription supports accessible solutions and alterations.

16.6 Conclusions

With the advance of technological possibilities for the average user and new kinds of content exploration, the visually impaired user should not be excluded. Given the new level of abstraction these new technologies permit, the possibility emerges to specify and implement a *fundamental* solution for this technological gap. The new goal changes from focusing on the effort to introduce *new accessible formats*, to ensuring the possibilities to merge the 'normal' content representation formats and their production/consumption tools with their accessible counterparts. This can be achieved by applying the Design for All paradigm at the design phase of new or updated products.

16.6.1 Design for All or Accessibility from Scratch

The implementation of the Design for All paradigm is about ingraining software and system design principles that are based on the findings of a multidisciplinary research and development approach towards the production and provision of accessible digital content. The multidisciplinary approach ensures sufficient 'openness' in the requirements met in the design. Given the transparent nature of the design and the multidisciplinary nature of these requirements, long-term effects of Design for All can be met. Examples of these long-term goals are education in accessibility and, *vice versa*, accessibility in education.

To ensure global acceptance of Design for All, a wider scope is needed which takes in more abstract notions. Given the level of specification, the Design for All notions can be implemented in a variety of applications. 'Big' formats, such as the MPEG initiative, have a big impact and a huge user base. Additionally, many software and hardware products are readily available and are therefore potentially accessible consumer products.

16.6.2 Applying Design for All in Emerging Standards

Global multimedia technology, such as the MPEG initiative, provides a transparent and modularised collection of knowledge and technology for the creation, representation, modelling and provision of multimedia content on a variety of hardware platforms. The system design of the individual MPEG family members provides primitive building blocks for an abstract accessible multimedia framework. Integration of accessibility services into the MPEG family, through the addition of digital item adaptation (DIA) (Jannach *et al.* 2005) definitions and additional components, could achieve huge advances in the availability of accessible information.

Expert knowledge on system design and applications of accessibility technology would be partly available in a widely used multimedia format. To be able to access this multimedia format, a widely available software development kit would be developed. The expenses of researching and developing personalised applications could be reduced. The technology *and* knowledge in an implemented, managed form would be available and this would create the possibility for a new range of accessibility products. Products that would enable the end-user with special accessibility needs to *create* (or deduce) accessible versions of multimedia content in their own time, in their own living environments.

16.6.3 Accessibility in Emerging Technology

Design for All and the accessibility issue in general most often seem to amount to a good understanding of generic programming and long-term design. Every technological evolution provides new opportunities and attributes. To be able to address these new opportunities and attributes, the model that governs the accessible services needs to be able to access these new opportunities. Without this adaptation the new features will remain hidden for the print impaired consumer.

Exactly the same phenomenon occurs with advances in non-accessibility related technology. Every now and then a new specification appears on the market and new and existing features are enhanced. For example, the video card industry has to invest a huge amount of funds and effort to 'activate' these new features. If the specification of a software driver architecture, including its development cycle, production schedule, logistic and knowledge management infrastructure, is not tuned towards this dynamic nature, the producers will be left behind, the consumers will be left behind and potentially technology will become inapplicable.

System Design and especially *educating* system design is about accessibility. How can collections of components be presented in coherent groups? How can these components be bound with communication structures that do not create a spaghetti structure? How can a transparent, scalable and manageable design be found for a very complex task? And ultimately, how can interoperability between all these components be ensured? These are all important topics in system design. Fundamentally this is the accessibility of knowledge at a fundamental level: knowledge describing the practice and theory of technology, science and its applications, and practical manifestations of knowledge, which is to say practical products. This is the wide scope of the Design for All paradigm.

Questions

Q.1 How many formats for music encoding have been proposed over the last 10 years?

Q.2 What are the main formats for the alternative presentation of music notation?

Q.3 Should we consider dyslexia as a print impairment or a cognitive impairment and to what extent does dyslexia present problems for reading music notation?

Q.4 What are the advantages and disadvantages of a three-layer information processing model?

Q.5 What are the main differences between *accessibility* and *usability* and at which points do they converge?

Q.6 To what extent have notions of accessible information processing been adopted within worldwide multimedia standards?

Projects

P.1 Structuring data

One of the key elements when building an accessible music system is analysing and processing the data, so that the structure can be re-used in alternative notation. Take a music score and try to list the music notation hierarchically. Is there more than one way of doing this? Are all hierarchies relevant, or is there a correct hierarchy? How does the meta-information about the piece (title, composer, global time and key markings) relate to this hierarchy?

P.2 Data description
Take a short musical piece and try to describe all the content (it may be helpful to use the hierarchy created in Section 7.2.1). Which information is the most important? Is there information that is optional for some users, such as fingering specific to certain instruments? Should this information be presented?

P.3 Accessible music services
Try to find out how visually impaired musicians are provided for in your country. How are the scores processed? Do the specialist organizations embrace technology and standards, or is it a manual process? How long on average does it take for a score to be produced? Write a report that analyses these findings and suggests where the Accessible Information Processing chain can be improved.

References

Crombie, D., Lenoir, R., and McKenzie, N., 2003, Producing Accessible Multimedia Music, Proceedings WedelMusic Conference 2003, Interdisciplinary Centre for Scientific Research in Music, University of Leeds, Leeds, UK

Jannach, D., Leopold, K., Timmerer, C., and Hellwagner, H., 2005, "Semantic Multimedia Adaptation Services for MPEG-21 Digital Item Adaptation", W3C Workshop on Frameworks for Semantics in Web Services, Innsbruck, Austria

Krolick, B., 1996, New International Manual of Braille Music Notation, World Blind Union, FNB, Amsterdam, Netherlands

Resources

http://homepages.cwi.nl/~media/www9/markku/pw-wap-01.html
Applying a Navigation Layer to Digital Talking Books: SMIL, XML and NCX

http://members.optusnet.com.au/~terryk/toccata.htm
The Toccata project

http://mpeg-21.itec.uni-klu.ac.at/
An introduction to MPEG-21 and DIA

http://msdn.microsoft.com/library/en-us/SAPI51sr/Whitepapers/srportingguide.asp
The Microsoft SAPI speech engine

http://www.borland.com/bes/
Borland application servers for J2EE™ platform applications and Web Services

http://www.chiariglione.org/mpeg/
More MPEG documentation under 'Hot news' or 'Working Documents'

http://www.cordis.lu/fp5/home.html
Homepage of the European Commission Community Research – the Fifth Framework Programme

http://www.daisy.org
The DAISY consortium

http://www.daisy.org/tools/playback.asp
DAISY software and hardware playback tools

http://www.daisy.org/z3986/default.asp
DAISY format DAISY 3.0/NISO z39.86 standard
http://www.dancingdots.com/
The GOODFEEL® Braille Music Translator
http://www.design-for-all.org/
EIDD, The European Institute for Design and Disability
http://www.dodiesis.com/asp/bmk.asp?language=1
The PLAY 2 project
http://www.dodiesis.com/asp/whatisBME.asp?language=2
The Braille Music Editor
http://www.duxburysystems.com/dbt.asp?product=DBT%20Win
The Duxbury Braille Translator
http://www.edean.org/, http://www.e-accessibility.org/
EdeAN, The European Design for All e-Accessibility Network
http://www.euain.org/
The European Accessible Information Network
http://www.ingentaconnect.com/search/expand?pub=infobike://mcb/238/2001/00000019/00000001/art00006&unc=
Worldwide training and technical support for DAISY
http://www.ingentaconnect.com/search/expand?pub=infobike://mcb/238/2001/00000019/00000001/art00002&unc=
DAISY Consortium: information technology for the world's blind and print-disabled population – past, present, and into the future
http://www.interactivemusicnetwork.org/
The Music Network project
http://www.interactivemusicnetwork.org/mpeg-ahg/
Music Network MPEG Ad Hoc Group on Symbolic Music Representation
http://www.interactivemusicnetwork.org/wg_notation/
Music Network Notation Workinggroup
http://www.lilypond.org/web/
The LilyPond automatic music notation system
http://www.midi.org/about-midi/specshome.shtml
Musical Instrument Digital Interface (MIDI) specification
http://www.mpeg.org
MPEG, the Moving Picture Experts Group
http://www.mpegif.org/
MPEG industry forum
http://www.musique.umontreal.ca/personnel/Belkin/NIFF.doc.html
Notation Interchange File Format (NIFF)
http://www.mutopiaproject.org/
The Mutopia Project: free sheet music for everyone
http://www.sightandsound.co.uk/pages/JAWS.htm
The JAWS screenreader software
http://www.s-line.de/homepages/gerd_castan/index_e.html
Common music notation and computers: NIFF, SMDL, DARMS and GUIDO
http://www.sun.com/software/product_categories/application_integration_messaging.htm
Sun Java System Application Server and integration services, a platform for the standards-based deployment of Java applications and web services
http://www.w3.org/TR/SVG/
Scalable Vector Graphics (SVG) 1.1 Specification
http://www.w3.org/XML/
Extensible Markup Language (XML)
http://www.wedelmusic.org/ Homepage for the WedelMusic project

17 Assistive Technology for Daily Living

Learning Objectives

The *activities* attribute of the comprehensive assistive technology (CAT) model, developed by the authors and presented in Chapter 1, has six main categories, of which three are categorized as fundamental activities and three as contextual activities. The assistive technology required to support blind and visually impaired people in carrying out activities in the fundamental activity categories, mobility and communication and access to information, has been discussed in previous chapters. Most blind and visually impaired people do not require assistive technology to support activities in the third category, cognitive activities.

This chapter and Chapter 18 will discuss the assistive technology required to support blind and visually impaired people in carrying out contextual activities. This chapter will focus on the technologies required for daily living, whereas Chapter 18 will consider the technologies required to support blind people in education, employment and recreational activities. The chapter will also consider the solution of accessibility issues associated with communications and information technologies other than computers and the Internet, telecommunications and print media. Despite being a component of one of the fundamental activities rather than a contextual activity, this access to information and communication topic fits logically into this chapter, since it includes the use of smart cards and information kiosks, which are mainly used in daily living applications.

As the name implies, daily living involves a wide range of different activities that are carried out on a regular basis, often every day or even several times a day. Most of these activities take place in the home environment, though some of them, such as shopping, take place outside it. The CAT model is used to provide the categorization of daily living activities required in order to provide a comprehensive treatment of the assistive technology needed to support these activities and enable blind and visually impaired people to live independently in their home setting. The assistive technologies used to overcome accessibility barriers in daily living range from simple low technology devices to very sophisticated and specialized high technology devices, giving the chapter a large number of subsections and coverage of a very wide range of different types of engineering solutions. The chapter provides descriptions of devices and, in some cases, details of the engineering principles involved.

The learning objectives for the chapter include the following:

- Using the CAT model as a framework for categorizing the assistive technology used in daily living.
- Obtaining an overview of the wide range of assistive technology solutions used in daily living applications.
- Understanding the engineering principles of some of the assistive technology solutions.

17.1 Introduction

This chapter provides an overview of some of the technologies that are used to aid blind and visually impaired people in performing everyday activities. On the one hand, a number of different, very useful, and often ingenious devices are available. On the other, many of these devices are quite expensive, often several times the cost of the equivalent devices for sighted people, and there is nothing like the range of choice available to sighted people. There is also very little information about the extent to which different devices are used and what blind and visually impaired people think of them. It is also possible that lack of information and high costs are preventing blind and visually impaired people obtaining all the devices they could benefit from. There are also issues of whether device design is always as good as it could be or whether some people are discouraged from using particular devices due to factors such as appearance, poor quality documentation or difficulty in using the device. In some application areas, assistive devices are not available or there are only a few products and devices that are designed specifically for blind and visually impaired people. On the other hand, there are a number of products and devices where vision is not required or which have been designed using good design principles, including Design for All, so they can be used by a wide range of people, including blind and visually impaired people.

Most devices are available with English language markings for controls, English speech for the audio interfaces and English documentation. A number of devices are available that also use a selection of European languages. Although there are probably additional devices in other languages about which little information is available in Europe, blind and visually impaired people who do not speak English and require devices with speech output or documentation in languages other than English could experience some problems.

The *activities* component of the CAT model presented in Chapter 1 is used to provide the structure for the presentation of the different assistive technology solutions considered in this chapter. Figure 17.1 shows the components of this third level branch of the model, which will be used to systematically detail the assistive technology solutions available. Thus, *daily living* involves a wide range of activities relating to personal care, care of the home environment and health. Most, but not all, of these activities are carried out in the home, but associated daily living activities, such as the use of money and shopping, are also included.

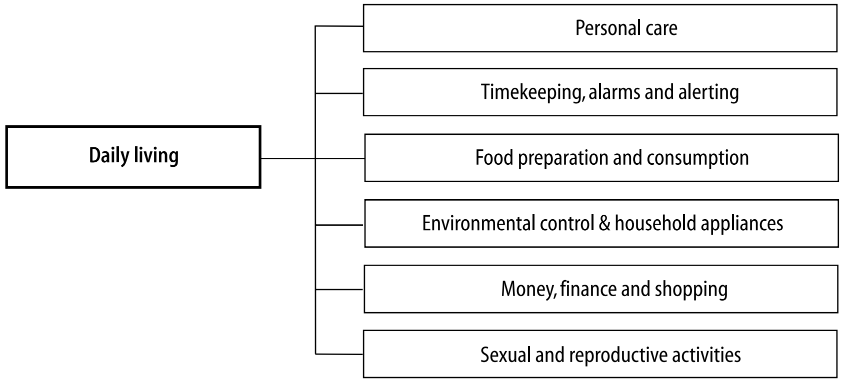


Figure 17.1. Activities in the area of *Daily Living*

However, it should be noted, that in a particular activity area only the assistive technology solutions relevant to visually impaired and blind people are given.

Author Note. It is in the nature of assistive technology for daily living that the authors will describe a wide range of products. These have been chosen to illustrate typical features that can be found in available products. The selection made is not comprehensive, merely illustrative. Readers seeking comprehensive lists of products are urged to consult a database like that of the ABLEDATA website or visit the websites of the many assistive technology distributors. A list of useful websites can be found in the Resources section at the end of the chapter.

17.2 Personal Care

In the activity area of *personal care*, blind and visually impaired people require assistive technology to identify the various items of their clothing, but not to get dressed or wash. Thus, this section considers two main topics: labelling systems, and health care monitoring and the use of medicines. Systems for tactile or audio labelling of items used in daily life, including clothing, are very important for blind and visually impaired people. However, it should be noted that the labelling systems described here have a number of other applications in addition to those relating to personal care. The use of vision in health care monitoring and the administration of medicines is very important and a range of assistive technology systems to assist blind and visually impaired people with these activities is described in this section.

17.2.1 Labelling Systems

Being able to distinguish objects is clearly very important. Most sighted people rely on vision. However, some information is given by touch and sometimes by smell and taste, though it is often not appropriate to try to distinguish objects by taste.

There is also the problem of discriminating between different items that come in similar containers and therefore feel very similar. Smell may also not be a reliable guide to the contents of these containers. Therefore, there is a need for devices that can be used to enable blind and visually impaired people to identify objects.

A variety of different approaches has been used. All of them involve a small label being attached to the object. The approaches can be divided into low-tech tactile and high-tech audio, though there are also some higher tech tactile approaches. In the low-tech approaches, the labels used either have a distinctive shape or are marked with Braille, moon or the Fishburne alphabet. The high-tech approaches use a number of different technologies including radio frequency identification (RFID) and bar code scanning.

17.2.1.1 Tactile Labelling Systems

Tactile labels using Braille

There are a number of different products for producing Braille labels. Say What is a clothing marker which fits over the hook of a hanger to identify the clothing on it. It consists of a strong plastic tab with a circular looped end and comes with sufficient tape to make 20 7.5 cm long labels with a Braille labeller (which is not included). Each kit includes 10 tags that can be used repeatedly. Do-Dots, model 30080, are clear clothing markers with Braille markings that can be fastened to the hem, cuff or collar of clothing. They consist of two button parts that snap together. On one side, they are marked with the Braille for the design, namely light or dark, print, plaid, stripes or plain, and on the other side use one of 45 Braille symbols for the colour. Each package of 100 buttons includes a Braille-coded key.

Braillables from American Thermoform Corporation can be cut to size and will stick to any smooth plastic, metal or painted wood surface. They can be used for book titles, labelling shelves, canned goods and other objects. They come in packets of 12 single 8.5 in × 11 in (21.6 cm × 27.9 cm) sheets. The automatic Braille labeller from KGS Corporation in Japan produces transparent, adhesive Braille labels. Characters, numerals and symbols are input and edited in the LCD display. The 10 function keys allow up to 10 words or phrases to be recorded and retrieved from memory. The length of the tape required to print the Braille label is automatically calculated and displayed on the LCD. Due to the use of a LCD display, using this device probably requires assistance from a sighted person.

The Braille Dymo gun from the RNIB embosses Braille letters onto Dymo tape. The alphabet dial has both Braille markings and black lettering on white, so it can be used by Braille and non-Braille readers. One 3-m roll of tape is supplied and additional rolls can be purchased in various widths. The Dymo gun produces Grade I Braille and some Grade II contractions. The RNIB also distributes plastic self-adhesive sheets for producing self-adhesive Braille labels on a Braille machine or hand frame and a plastic faced magnetic rubber strip which can be Brailled on a hand frame with firm pressure. It can be stuck to metal objects, unless they have high aluminium content.

Tactile Colour from Tactile Colour Ltd consists of a pack of self-adhesive coloured vinyl sheets, with each colour having a distinct texture. Shaped pieces can be cut from the sheets and used to mark equipment. The labels can be distinguished by texture, colour or shape. They could also be used to indicate the colours of items.

Tactile labels using the Moon and Fishburne alphabets

The Moon alphabet is a tactile alphabet which is simpler for older people to learn than Braille. It is mainly used in the UK. The RNIB produces various materials for labelling items with Moon. Moon alphabet sheets contain moon letters, punctuation marks, numerals and other signs which can be used to label CDs, food containers and other items. The characters are available in 24-, 36- and 48-point size. The moon hand frame can be used to produce moon characters on plastic film or self-adhesive labels. The film is placed between the hinged layers of the frame and each Moon character is drawn on a square in the grid. A pen and 100 sheets of A5 plastic film are supplied. The characters can be felt on the film when it is removed from the frame and placed on a solid surface. The Fishburne alphabet is a tactile alphabet designed to provide simple touch identification for household and personal items. It does not replace Braille, but may be a good alternative for people who have reduced sensitivity in their hands, possibly as a result of diabetes.

17.2.1.2 Radio Frequency Labelling Systems

Radio frequency identification (RFID) principles and issues

Radio frequency identification (RFID) carries out automatic identification by storing and remotely retrieving data using RFID tags or transponders. The tags contain silicon chips and antennae to enable them to receive and respond to radio frequency queries from an RFID transceiver. Passive tags do not require an internal power source whereas active tags do. The very small electrical current induced in the antenna by the incoming radio frequency signal provides sufficient power for the CMOS (complementary-symmetry/metal-oxide semiconductor) integrated circuit in the tag to transmit a response. Therefore, the antenna (aerial) has to be designed so that it can collect power from the incoming signal as well as transmit the outbound backscatter signal.

The tag chip contains EEPROM (electrically erasable programmable read-only memory) for storing data. The absence of a power supply reduces the size of the device and there are even commercially available products that can be embedded under the skin. The smallest commercially available devices in 2005 were 0.4×0.4 mm and thinner than a sheet of paper. There are passive, semi-passive and active tags. Passive tags have reading distances of about 2 mm to a few metres depending on the chosen radio frequency. Semi-passive RFID tags are similar to passive tags, but they have a small battery. This allows the tag integrated circuit (IC) to be constantly powered and means the aerial does not need to collect power from the incoming signal and can therefore be optimised for the backscattering signal. Semi-passive RFID tags have a faster response and are therefore more effective in reading radio than passive tags (Anon 2006a).

Active RFID tags or beacons have their own internal power source which both powers their ICs and generates the outgoing signal. They often have longer ranges and larger memories than passive tags and can store additional information sent by the transceiver. Many beacons operate at fixed intervals to reduce power consumption. The smallest active tags are the size of a coin. They have practical ranges of tens of metres and a battery life of up to 10 years.

The majority of RFID tags are passive, are inexpensive, with an average cost of 0.20 Euro at high volumes, and do not require a battery. As universal RFID tagging of individual products becomes commercially viable at very large volumes, the cost will reduce further. Successful performance of close to 100% is required. Despite the cost advantages of passive tags, accuracy and improved performance in the presence of water and metal and improved reliability means that there are still many applications for active tags. RFID can be divided into four main frequency bands categorised by their radio frequencies: low frequency tags (125 or 134.2 kHz), high frequency tags (13.45 MHz), UHF tags (868–956 MHz or 463 MHz) and microwave tags (2.45 or 5.8 GHz). Using UHF tags globally requires them to be specially tailored to regional regulations, as there are no unified regulations for this radio frequency band.

The components of RFID system generally include tags, a tag reader, edge server, middleware and applications software. Each tag has a transponder which contains a digital memory chip with a unique electronic product code. The antenna has a transceiver and decoder. It emits a signal which activates the RFID tag to enable the signal to read and write data to the tag. The data transmitted to the tag could include identification or location information and/or details about the product tagged, such as price, colour and size. When an RFID tag passes through the RFID reader's electromagnetic zone, it detects the reader's activation signal. The reader then decodes the data encoded in the tag and passes it to the host computer, where the data is processed, often using Physical Markup Language (PML).

However, though relatively inexpensive, RIFD tags are still more expensive than bar codes and therefore not likely to totally replace them in the immediate future. A more serious concern relates to privacy issues. RFID tags on products continue to function after a product has been purchased and could therefore be used for surveillance. Although RFID tags are intended for short distance use, they can be scanned at a distance by a high gain antenna without the knowledge of the owner of the item. Purchasers will not necessarily be informed of the presence of a tag or how to remove it. If tagged items are paid for using a credit or loyalty card, then the tag identity could be tied to the purchaser's identity, giving a potential for abuse (Anon 2006a).

In the US the Department of Veteran Affairs Outpatients pharmacies are now putting 13.56-MHz talking prescription tags on medicines for visually impaired veterans. The information stored on the tags includes the name of the medicine, instructions for taking it and any warnings. It can be read by a battery powered, talking prescription reader.

RFID talking labels

The SHERLOCK talking labeller has a capacity of 32 MB, giving about 2000 labels or 2.5 h of spoken text. The speech output has three different volume levels. The voice recordings use RFID technology. The device is light and easily portable and works from two AA (rechargeable) batteries, with a set of batteries lasting for several thousand measurements.Twenty-five adhesive labels and ten non-adhesive labels are provided.

Tag It from Dräger and Lienert is an identification and search system, which uses RFID technology. Stamp-sized stick-on tags containing a coil and a small chip are taken from a dispenser and stuck onto objects. The RFID reader identifies the object by computer and speaks the description in a synthesised or natural voice. Tags can be recorded using the computer keyboard. To find an item the user types in the name of the item, for instance, 'dictaphone'. Tag It provides a list of the dictaphones in the system and the user chooses the one they want. The RFID reader then searches for this type of Dictaphone and informs the user when it has been found.

Other talking labels

The VOXCOM 50/100 label reader from Cobolt Systems can label items with cards on which messages of up to 10 s are recorded. A card is inserted into the device, the 'record' button held down and the message spoken. The message will be read back when the card is subsequently inserted into the device. Cards can be re-recorded as often as desired. The cards are plastic and waterproof and could be used to label products in a freezer. They can be attached with rubber bands, velcro, magnets and cable ties.

The VoiLa! talking bar code reader from ASSISTec Ltd is a hand-held wireless pen-style device, which uses three buttons to record messages to describe objects. It is wiped along the label which is moulded to act as a guide for the reader. The user then speaks the name they wish to give the item. To identify the object subsequently the reader is wiped along the label and the previously recorded name is spoken. Names can be changed or deleted, allowing labels to be reused. The VoiLa! can record dosage and times for medicines. It has an extended memory for long recording time, a low battery alert and battery failure does not cause memory loss. The device is supplied with 50 self-adhesive labels and additional labels are available in sheets of 50. A maximum of 250 different labels can be made with a recording time of 2.3 s per label. Multiple copies of the same label can be used, for instance to label several tins of peas. The device is powered by two AA batteries and instructions are provided on cassette.

Talking Labels from the RNIB can be used to identify a variety of household items. The labels have a clip so they can be attached to products. Up to 60 s of message can be recorded by pressing the record button on the back of the recording device with a pen and speaking into it. The message can be replayed and a new message recorded over the previous one. Talking tins from the RNIB can be used to identify the contents of tinned food, storage containers, bottles and sprays.

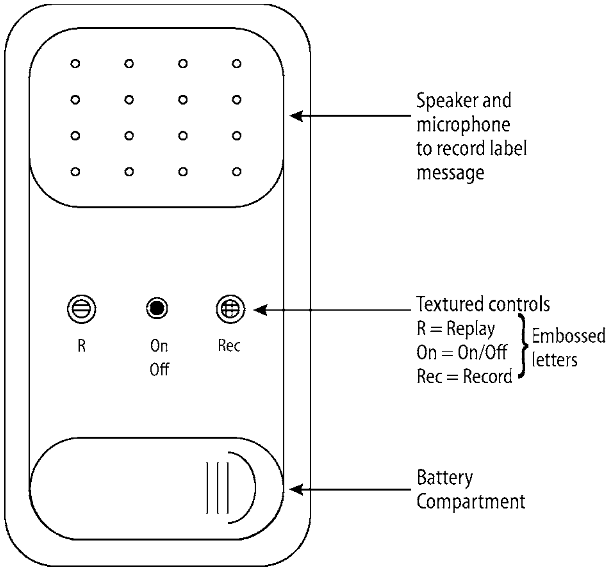


Figure 17.2. Sketch of a typical talking label device

It fits on all standard tin lids, including ring pulls. The built-in magnets secure the talking tin to any steel containers and the self-adhesive metal disc or elastic strap provided can be used to secure the talking tin to other items. Descriptions of up to 8 s can be recorded. The Recordable voice pad can record a message of up to 10 s which can be added to greetings cards or gifts to make them more personal or used to identify items.

The sketch in Figure 17.2 illustrates typical features of talking label devices.

17.2.1.3 Barcode Labelling Systems

Talking bar code readers that make contact with the label

Bar codes comprise a series of printed horizontal strips of various widths with the digits zero to nine represented by different bar patterns that can be read by a laser scanner. Retailers use bar codes to record prices and generally have extensive databases of barcodes. A typical barcode is shown in Figure 17.3.

6 87524 61062 7

Figure 17.3. Example of a UPC bar code

Bar code readers generally comprise the following three main components:

- The bar code scanning module, which inputs the bar code from labelled items and transfers it to the database application *via* laser beam or contact type scanners. In the case of the ScanTell the scanning module accepts universal product code (UPC), code 39 and other bar codes and transfers the data in the form of ASCII code.
- The database application, which extracts item information from the relational database. Some talking bar code readers have both existing databases and the capacity to enter new items. The database can generally be used to edit bar codes.
- The voice synthesiser, which provides audible information using item data from the database. It also provides spoken prompts, guidance, notifications and warnings to users.

There are a number of different bar code readers with slightly different features. BarTalk developed by Kentucky Department for the Blind allows users to scan bar codes and hear the corresponding descriptions spoken by a text-to-speech synthesizer. The BarTalk software is a menu driven program that is distributed free as shareware and does not require screenreader software. The BarTalk scanner is a Metrologic SMDE700i laser bar code protection scanner housed in the same box as a Walk ‘n Talk speech synthesiser with a power supply, connections and a volume/on/off switch. The BarTalk scanner is connected to the serial port of an IBM compatible PC. Users who already have a speech synthesiser only need to connect an SMDE700i laser bar code protection scanner. Since the software was developed several years ago, operating system requirements are stated in terms of compatibility with DOS 2.0 and above.

Scan and Say is a talking barcode program for Windows from the RVB Systems Group. Its database already contains about 500,000 items and new items can easily be added by users. The system can easily be customised, with product descriptions determined by the user. For instance, cooking instructions can be added to the food descriptions and the system can be used to create and read aloud a shopping list. It can be used for a wide range of different types of items. A pre-printed barcode, of which 500 are provided with the system, can be added to an item that does not already have one. Scan and Say can also be used to time stamp items, which is useful in keeping track of when perishable food is out of date and the last time a medication was taken. A barcode scanner is provided with the system. It can be attached to a USB, RS232 or keyboard port with a 15-ft (4.6-m) cable. It can be used in hand-free mode on the adjustable stand or hand-held. There is a wide pattern of 20 scan lines rather than a single scan line.

ScanTell is a talking bar code reader that identifies and speaks the names of barcoded products. The code on the barcode label is read *via* a bar code scanner and the item is then identified from a cross-referenced relational database. The description of the item is output using voice synthesiser software and hardware. The system can operate on desktop and portable PCs. It uses audible user prompts and minimal keyboard input. It has a database management system with user-definable databases. ScanTell is available for Windows and DOS. The voice synthesiser can be run from either the hard drive or CD-ROM if space on the hard drive is limited, but is faster on the hard drive. The system is provided in a modular form, so that users only need to purchase the components they need.

The Scanacan for Windows Deluxe, Scanacan Professional Deluxe and Scanacan Professional Elite from Ferguson Enterprises are talking bar code readers. All the systems have databases of 84,000 grocery items and 62,000 hardware items. Users can scan items into the databases by passing the bar code scanner close to the item's bar code. They can also create new databases. The systems include Scanacan software on CD-ROM, the two databases, a JAWS for Window script set and a Metrologic Omnidirectional Laser Bar Code Scanner, a strut and bracket for wall mounting and 100 pre-printed bar code labels. Instructions are provided in print, on CD-ROM and on cassette. Scanacan for Windows is intended for home use, whereas the other two systems can be used by blind and visually impaired store owners. The Scanacan Professional software provides support for inventory and point-of-sales operations and the Professional Elite system includes an Epson cash drawer and sales-receipt printer. The system can be used on IBM compatible computers and the cash drawer, receipt printer, pre-printer bar code labels and clothing labels are also available separately.

Remote scanning of bar codes

The ID Mate from Envision America is a talking bar code identifier that can read a bar code from a short distance away. It consists of a recorder and an omnidirectional scanner with a small screen for scanning connected by a cord. The ID mate identifies an item by scanning a bar code and allows the user to record a description. When the same type of item is subsequently scanned the description will be spoken. Since a product database is not included, assistance from a sighted person is required to identify objects for the first time. There would therefore be additional benefits in including manufacturers' databases of information to enable products to be identified immediately. A number of blank bar codes that can be attached to items are included for items that do not have a bar code. Bar codes are identified relatively quickly for most items, though problems can be experienced with bar codes on dented areas of cans and boxes. The ID Mate is used by some blind people who manage vending stands, but is rather expensive for purely home use.

It has a memo function, which can be used to record reminders, such as phone numbers and appointments. The system is provided with an earphone, microphone jack, and a 4 MB flash memory card that can store 420 five second bar code messages or 35 min of digital recording. It can also use 32 MB cards. The manual is provided in medium print and audiocassette versions and can also be downloaded. A shorter version of the manual can be accessed in the 'help' mode.

Typical features of one type of remote barcode scanning system can be seen in the sketch in Figure 17.4.

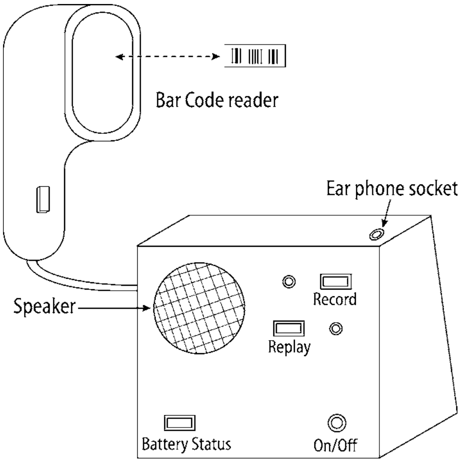


Figure 17.4. Sketch of one type of remote barcode scanning system

17.2.2 Healthcare Monitoring

To facilitate independent monitoring of personal health and wellbeing there are a number of devices with audio and/or tactile output for measuring temperature, blood pressure and body weight. Being able to read the labels on medicines is also an important aspect of personal (health)care.

17.2.2.1 Clinical Thermometer

The Talking Ear Thermometer from Cobolt Systems has a probe that is inserted into the ear. Thirty readings can be stored and recalled at any time. The device can also be used as a talking clock with a calendar display. Two AAA batteries power this particular thermometer. MAB Community Services' talking clinical thermometer has a digital display and a fast response time. It remembers the last measured temperature. The RNIB talking digital body thermometer reports the temperature reading and automatically switches the power off after 7 minutes. The last measured temperature is stored in the memory. It uses one 3 V lithium battery.

17.2.2.2 Blood Pressure

Combined talking blood pressure monitors and pulse measuring devices are available from both Caretec (called the SweetHeart) and Cobolt Systems. Both devices are operated by a single button, which automatically inflates the upper arm cuff, takes measurements, announces the measurements and deflates the cuff. Values of pulse rate and systolic and diastolic blood pressure are displayed. The Cobolt Systems device uses fuzzy logic techniques to ensure that the cuff inflates in proportion to the blood pressure. The previous eight measurements can be displayed and the last one spoken. The SweetHeart shows measurements on the LCD display

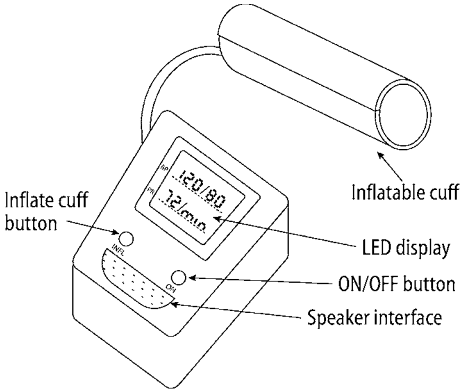


Figure 17.5. Typical features of talking blood pressure meters

in large letters and announces them several times in a choice of 25 languages. The measurements can also be shown on a Braille display. Its blood pressure measurements are claimed to be accurate to ± 3 mm Hg and pulse measurements to $\pm 3\%$. Typical features of accessible blood pressure meters are shown in Figure 17.5.

Cobolt Systems also distribute a small blood pressure and pulse monitor that straps to the inside of the wrist. The device is switched on at the press of a button and then inflates the wrist cuff, announces the blood pressure, deflates the cuff and automatically switches off. The device is able to record sets of 99 values. The test results, date and time are automatically stored after each test and can be accessed and announced.

17.2.2.3 Blood Glucose Level

The SensoCard Plus meter automatically stores the last 150 blood sugar test results and stamps the data with the time and date of the reading. Either the press of a button or inserting a test strip will switch on the meter so that it is ready for reading. The device is provided with a lancing device, eight disposable lancets and a check strip. Packs of test strips can be obtained from Cobolt Systems or sometimes on prescription. The meter has full UK medical approvals. The optional LiteLink can be connected to a computer RS232 serial port, allowing data to be transferred from the meter to the computer.

The AccuCheck Voicemate from Roche Diagnostics is a voice synthesiser system incorporating the AccuCheck Advantage monitor and an audible insulin vial identifier (for Lilly insulin only). It uses Comfort Curve test strips. The use of a code key system eliminates the need for sighted assistance in indexing the meter to a new batch of test strips. Large print and audiocassette instructions are available in English and Spanish. The AccuCheck Advantage monitor provides a large LCD readout.

17.2.2.4 Insulin Measurement and Delivery Devices

The Palco Load-Matic from Diabetesstore.com is a tactile insulin-measuring device that accepts BD 100-unit syringes and aligns the needle with the vial stopper. There are two separate controls for single and ten unit increments respectively. There is a tactile prompt to confirm the dose setting. The Count-A-Dose Insulin Measuring Device from Medicool is a gauge calibrated for use with U-100 vials and BD 0.5-cc (low dose) syringes only. When a thumb wheel is turned clicks are heard and felt for each one unit increment measured. The device can hold one or two vials of insulin for mixing and the needle penetrates the vial stopper automatically. Instructions are provided in print and on audiocassette.

17.2.2.5 Talking Bathroom Scales

Cobolt Systems distribute talking bathroom scales. These scales are available in a number of different European language versions. They use advanced strain gauge technology and electronic calibration. All language versions give weights in kilos and some versions give weights in other units as well. The English version gives weights in kilos, stones and pounds, or pounds and can convert between these units. The American and Spanish versions can give weights in pounds and kilos. A memory function in the bathroom scales will record the weight and state how it has varied from the last weighing. Four separate memory buttons allow different individuals to weigh themselves and monitor their weights individually. The unit also displays the weight in large numerical digits.

CareTec distributes talking bathroom scales manufactured by Soehnle with a large LCD display, speech output in 16 different languages, automatic switch off and the facility to weigh in metric or imperial units. These talking bathroom scales weigh in intervals of 100 g or 10 oz (ounces).

17.2.2.6 Braille and Audio Labelling of Medicines

The European Commission has issued a directive, Guidance concerning the Braille requirements for labelling and the package leaflet (EC 2005), but it currently only applies to medicinal products approved after 30 October 2005. The directive does not state which layer of the packaging should have the Braille label.

The ScripTalk System from En-Vision America provides voice output from an electronic speech synthesizer. The Talking RX Audio Label from Millenium Compliance Corporation is a voice recording system where each 'label' is also a digital recorder/player. People using several different medicines will require several Talking RX devices, which are fortunately relatively inexpensive. Each 'talking label' can be repeatedly reprogrammed. The MedGlider is a combination pill organiser and talking timer. Up to four daily alarms can be set for the exact times to take medication. The alarm can either beep or say 'time to take your medicine'. There are two versions, System 1, which includes one pill case and the talking medication timer, and System 7, which includes a storage tray, seven individual pill cases and the talking medication timer.

17.3 Time-keeping, Alarms and Alerting

Time-keeping is very important in the modern world. Being late tends to have negative consequences, including missed appointments and arriving after the shops or the library have closed. Timers and stopwatches are also used in cooking, sports and other applications.

In addition to time-keeping devices, various alarm and alerting devices are described in the second part of this section. These are designed to make visually impaired and blind people aware that a critical safety condition has occurred and that they may need to take some action, such as calling the fire brigade and/or evacuating their home in the case of a fire.

17.3.1 Time-keeping

Most time-keeping devices, namely, clocks, watches and timers have visual displays, though they provide audible alarms and alerts. There are a number of clocks, watches, timers and multifunction time devices for blind and visually impaired people. Most of them have speech output, but there are also timepieces, and timers with Braille or other tactile displays.

17.3.1.1 Manual Tactile Clocks

The Touch-Time Desk Clock is a wind up clock with a 5 cm (2 in) face with the hours marked by tactual dots. Three dots mark the 12 o'clock position, two dots mark the 3, 6 and 9 o'clock positions and a single dot each of the other numbers.

17.3.1.2 Electronic Clocks with Speech Output

There are a number of different talking alarm clocks. Many of them have additional features such as calendar, timer, stopwatch and/or radio. The ZeitGeist from Caretec is a talking clock with a variety of calendar, scheduling, stopwatch, timer and alarm and alerting functions. It provides speech output with a choice of 20 different languages and three volume levels through an earplug or a loudspeaker. It is very light and can be carried in a pocket or worn round the neck.

Talking alarm clocks are easy to purchase. For example, Cobolt Systems distribute several talking alarm clocks. The T-12 talking alarm clock speaks the time and data. It has four alarm settings over a 24 h period, as well as an optional hourly chime and a stopwatch function. The 'Atomic' talking radio controlled calendar alarm clock is synchronized to the Rugby time signal, making it very accurate and it does not require setting or adjusting. If the clock is unable to receive the radio time signal, then it functions as a normal quartz clock and can be set manually. It has a large LCD display and an optional hourly time report. The Rugby time signal is broadcast by the National Physical Laboratory from the very low frequency transmitter near Rugby. It can generally only be received in the UK. The effective radiated power of this signal is 15 kW at a frequency of 60 kHz. The signal is transmitted at one bit per second, with long (200 ms) and short (100 ms) pulses representing the binary digits '1' and '0' respectively. The entire signal consists of 43 bits and starts on the seventeenth second of each minute. Consumer clocks are generally updated once an hour from this signal. On 1 April 2007 National Physical Laboratory time signal transmissions will officially switch to a transmitter at Anthorn, in Cumbria and the Rugby signal will be switched off.

The 3-Event Talking/Recording alarm clock can record reminders of up to three separate events in the user's own voice. Each recording can be up to 6 s long and can be used as an alarm. The large LCD display shows the current date, the event date(s) and event count-up/down times. A reminder function can also be used as a daily wake-up alarm. The current time is announced at the press of a button. Cobolt Systems also distributes an Urdu talking alarm clock with separately adjustable speech and alarm volumes and a choice of alarm sounds. The current settings are reported at the touch of a button and there is an optional hourly report. The time and settings are not lost when changing the battery.

Independent Living Aids distribute several different talking clocks. The Seiko talking disk clock has both a digital readout display and provides a spoken time at the press of a button, with volume control, three different alarm sounds, a snooze option and an hourly chime function. The Teeny Tiny Talking Alarm Clock and Radio is a very small talking clock radio, with a large LCD clock display, that plays FM stations. The clock talks the user through the setting function and the alarm has the sound of a rooster crowing. The Can-Can Talking Cube Clock can be set to a whisper or almost shout. There are three alarm sounds and the alarm will sound for 1 min and then repeat at five 4.25 min intervals unless it is turned off.

Talking wristwatches of various levels of sophistication with prices that are comparable with those of non-talking wristwatches are widely available. Some versions have additional functions such as an optional hourly chime, alarm, calendar and stopwatch. In many of the watches, the setting buttons are recessed into the sides to prevent unintentional use.

17.3.1.3 Tissot Vibration Technology (Silen-T)

The Silen-T watch from Tissot uses touch-screen technology to enable users to tell the time through silent vibrations (see Figure 17.6). Solid and pulsing vibrations indicate the hour and minute settings respectively. The standard tactile watch reference points are indicated by raised visible markings around the bezel to enable both blind and sighted people to tell the time. The watch has a silent vibrating alarm. The case is water resistant to 30 m. There is a two-year power reserve and the touch sensitive crystal is made of scratch resistant sapphire with a non-reflective coating.

17.3.1.4 Braille Clocks

The Braille Digital Clock Calendar designed by the Kentucky Department for the Blind displays the time or the date on a six cell refreshable Braille display.

Figure 17.6. Silen-T watch (photograph reproduced by kind courtesy of Tissot S.A., Switzerland)

An audible alarm can be set to sound at any time during the day. The calendar is automatically adjusted between summer and winter settings and for leap years. The Braille display requires an adaptor to be powered from the mains. Backup batteries with a life expectancy of 10 years allow the internal clock to run continuously whether or not the device is plugged in. The clock accuracy is at least ± 1 min per month. The separate set and view modes reduce the likelihood that users will accidentally change clock settings. In view mode each press of the select button displays sequentially the current function (time, alarm and date), AM/PM status, alarm on/off status and the day of the week. All six Braille cells can be set one at a time in the set mode to give the desired time, date and alarm time. Setting the clock to the exact second is facilitated by zeroing the rightmost digit in the set time mode using the change button.

17.3.2 Alarms and Alerting

In this section, three different alerting devices are described, one is a device for timing the operation of household appliances, the second is an alerting alarm for smoke and carbon monoxide detection and the third is a freezer alarm. The section closes with a description of a class of alerting devices that can be used as a object finder or locator.

17.3.2.1 Talking Appliance Timer

Talking timers are available from Cobolt Systems, Dynamic-Living.com and the Ultmost Technology Corporation. They all function as count-up and countdown timers with a maximum time period of 24 h (minus 1 s) and also provide a talking clock. The Ultmost Technology device provides an automatic repeat of the count-down, which can be useful when taking medicines. The Dynamic-Living device is pocket sized, which makes it suitable for a wide range of applications, including exercise routines and team games, as well as a cooking timer and a reminder for

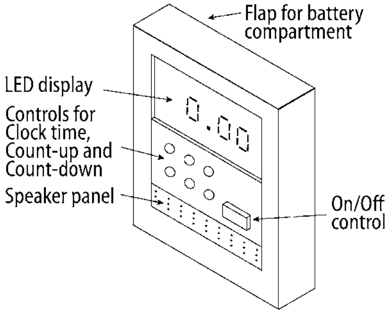


Figure 17.7. Typical features of talking timers

appointments or taking medicines. Its belt clip has a magnet for attachment to a refrigerator as well as a stand.

A talking timestat in English, French, German and Italian is also available. It has independent timers, which can be adjusted individually for each day of the week, a thermostat and a frost meter. Each of these functions can be used independently or in combination. The timestat can be used to control a mains operated appliance with a maximum rating of 2 kW, including electric heaters, tape recorders, radios, lights and electric blankets. When plugged into a standard socket the timestat charges the built-in battery. It remembers its setting for several months when disconnected or in the event of a power failure and can also function as a talking clock and thermometer.

Typical features of talking timers are shown in the sketch of Figure 17.7.

17.3.2.2 Talking Smoke and Carbon Monoxide Detector and Alarms

A typical talking smoke and carbon monoxide detector, for example, that distributed by Cobolt Systems, has a voice alarm which says 'fire fire' or 'warning carbon monoxide' with a very loud 85 dB alarm. It has a test/reset button and flashing LED to verify the detector operation. It speaks a warning if the batteries are becoming low. It is battery run and does not require any wiring.

17.3.2.3 Freezer Alarm

Cobolt Systems' freezer alarm is intended to be compatible with all freezers. It has a probe cable of about 1.5 m length, allowing the alarm to be placed at any convenient location outside the freezer. It gives a loud audible alarm and has a LED visual indicator to prevent accidental de-frosting of the freezer if the freezer door is left ajar or the freezer or power fails. The sketch in Figure 17.8 shows a typical freezer alarm device.

17.3.2.4 Object Locator

The Personal Locator System is based on the Talking Signs system (see Chapter 10). It comprises a transmitter that sends out a message on a beam of infrared light and a hand-held receiver. The transmitter is small and lightweight and battery powered, so that it is portable. The transmitter emits a beep sound when the receiver is close by. It has a sleep mode of operation to reduce power consumption when the transmitter is not activated by being searched for by the receiver. The locator system could be attached to various different objects, such as luggage or a seat on a train reserved by a visually impaired person, to help in finding the desired target object. A block diagram of the system is show in Figure 17.9.

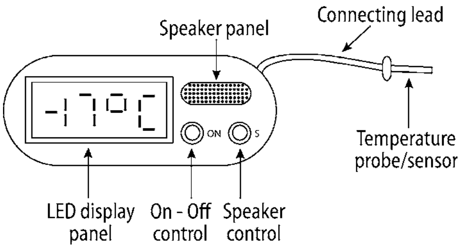


Figure 17.8. Sketch of a typical freezer alarm device

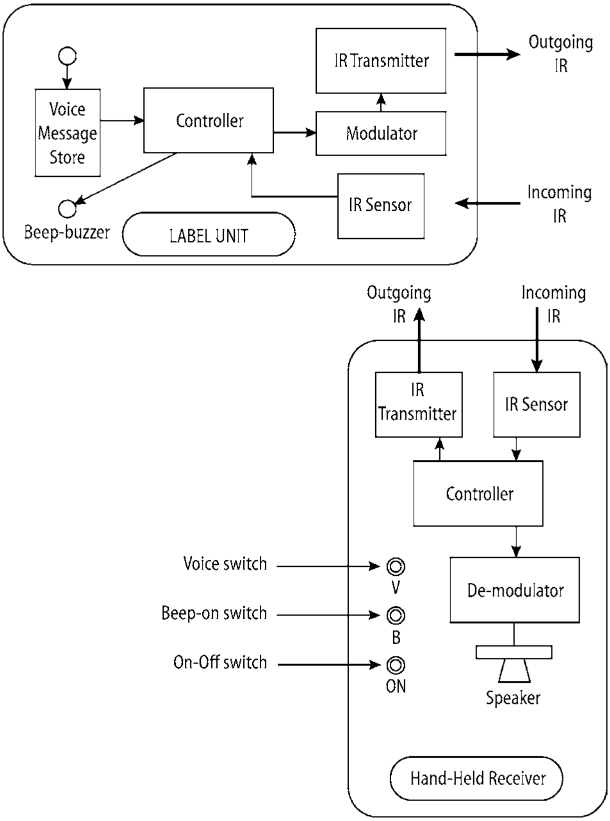


Figure 17.9. Block diagram of Personal Locator System

Independent Living Aids distributes a number of locators. The key locator whistle and flashlight is a small key chain unit that responds to the user whistling by beeping. It can be used to find items such as keys, remote controls, mobile phones and glasses (spectacles). It has a mini flashlight and a range of 10 ft (3.0 m). The 'now you can find it!' has four small locator discs that can be attached to items with the ring provided, tape or velcro. Pushing the appropriate button on the base of the locator will make the locator disc beep to guide the user to the item. The discs and corresponding buttons are colour coded and have Braille-like bumps to help users distinguish them. The range is 30 ft (9.1 m). There is also a version with eight radio frequency locators which both beep and flash when the corresponding button on the locator is pressed. The sound changes when the user gets within 40 ft (12.2 m) of the locator. The Super Sound Locator is a wireless transmitter/receiver with a very loud sound and a range of 100 ft (30.5 m). It continues making a noise until the stop button on the receiver is pressed. It can be used in baseball and basketball or as a pager. The National Federation for the Blind distributes a luggage locator transmitter and receiver. The receiver can be attached to any key chain and the transmitter to any luggage handle.

17.4 Food Preparation and Consumption

Visually impaired and blind people may experience a number of difficulties in the kitchen and with food preparation and consumption. Many of the technological solutions used in these activities are simple design modifications of crockery, cutlery and other kitchen tools to make them safe to use and to prevent food portions going astray. A typical example is a sheathed knife with a thickness guide to enable bread to be cut safely to a desired thickness. Another example is the use of rimmed plates and dishes and suckers on crockery to ensure that food remains in the right place during its consumption. Among the higher technological solutions in the kitchen are talking kitchen scales, measuring jugs, liquid level indicators, microwave ovens and food thermometers. Examples of these devices are described in the following sections.

17.4.1 Talking Kitchen Scales

Cobolt Systems distributes a talking kitchen scale in a number of different European language versions. It can weigh up to 5 kg or 11 lb in 5-g or 0.5-oz steps with a claimed accuracy of 2%. It has a large secure bowl and non-slip feet. The English-speaking version can convert weights between imperial and metric units. Touching the bowl gives a repetition of the weight. High and low tones sound every 5 g or 0.5 oz when the weight is adjusted slowly to help with precise measurements. There are adjustable volume and zeroing functions to allow a second substance to be weighed.

CareTec distributes talking kitchen and diet scales manufactured by Soehnle with a large LCD display with speech output in 16 different languages, automatic switch off and the facility to weigh in metric or imperial units. The kitchen scale has a Tara function which allows the scales to be zeroed without being emptied. Weights are measured in 2-g or 0.2-oz steps up to 2.5 kg or 5 lb 7 oz and in steps of 5 g or 0.5 oz up to a maximum of 5 kg or 11 lb. The diet scale has a Tara function and weigh in 0.5-g or 0.05-oz steps up to 1 kg or 2 lb 3 oz and in 1-g or 0.1-oz steps up to 2 kg or 4 lb 6 oz.

The tactile 2-lb scales from Science Products for the Blind have raised dot scales and can be used for weighing food, parcels and letters to calculate postage and other applications. There is a double dot every 4 oz (113 g) and single dots at every 1 oz (28 g) in between. They are accurate up to 2 lb (0.91 kg) and the zero adjust allows for the use of a container when weighing.

17.4.2 Talking Measuring Jug

Cobolt Systems also distributes a talking measuring jug (see Figure 17.10). It is calibrated for metric (litres and millilitres), imperial (pints and fluid ounces) and US (pints and fluid ounces) units and can convert its readings between these units. The user selects the type of liquid to be measured. The weight of the jug contents is then stated by synthesised speech at the press of a button or automatically as the liquid is added. The reading can be set to zero without emptying the container to allow a measured volume of a second liquid to be added. The jug can also be used for solids such as sugar. The container holds up to two litres of liquid. The speech volume is adjustable.

The RNIB distributes a 1$\frac{1}{2}$ l pint plastic measuring jug with a tactile gauge for the measurement of fluid cooking ingredients. There are standard print markings on the outside of the jug. The detachable gauge is placed inside the jug and held in place by a clip located at the top of the handle. The gauge has tactile metric markings

Figure 17.10. Talking measuring jug (photograph reproduced by kind permission of Cobolt Systems Ltd., UK)

on one side and imperial markings on the other side, so that measurements can be made in both types of units.

17.4.3 Liquid Level Indicator

Liquid level indicators use a sensor to detect the level of liquid in the container and give an audible indication when the level reaches a given distance from the top. Some of the indicators can be used with two liquids, but currently available devices do not give users the option to set either the relative volume of the two liquids or the desired level of the first liquid. There are also talking and tactile measuring jugs. Both the Sensa cup level indicator mk 3 from Cobolt Systems and the 3 Prong liquid level indicator from Lighthouse store can be used with one or two liquids. The Sensa indicator can be used with one hand and automatically adjusts itself for different sizes and types of container. It produces different audible indications when the liquid level is 0.5 in (1.27 cm) from the bottom and 0.5 in (1.27 cm) from the top. It is powered by two watch batteries, which the manufacturers suggest should last for the lifetime of the product. The 3 Prong indicator produces an intermittent beep at the lower level and a constant one at the top. The unit also vibrates. Science Products for the Blind's audible liquid level indicator has two small electrodes which act as liquid sensors. The speaker produces tones when the liquid reaches the electrodes. There are 10-cm and 30-cm models for home and laboratory use respectively.

17.4.4 Talking Microwave Oven

The talking microwave oven Mk 3 from Cobolt Systems and the LG talking microwave both have a high power output of 900 W and provide spoken confirmation of the functions selected and cooking time and announcements when the door is open and closed. They also both have a talking clock, talking kitchen timer and adjustable audio volume. The Cobolt device has a tactile wipe clean speech pad and 6 of the 12 control buttons of the LG device are embossed in Braille. The Cobolt device provides both basic and programmed operations and spoken requests to stir or turn food during cooking and to leave food to stand after cooking. Its instructions are available in large print and on tape. The LG device states the function of each button when it is pressed and has an '*incorrect key*' announcement when a button is pressed that cannot be used in a given mode.

17.4.5 Talking Kitchen and Remote Thermometers

Cobolt Systems' talking cooking thermometer has two temperature-reading cooking utensils—a detachable spatula and a fork. The temperature is shown on the large backlit liquid crystal display (LCD) and the degree of cooking is announced. The device speaks in English, Spanish and French. The fork and spatula are dishwasher safe, but the probe should be cleaned with a wet cloth and soap.

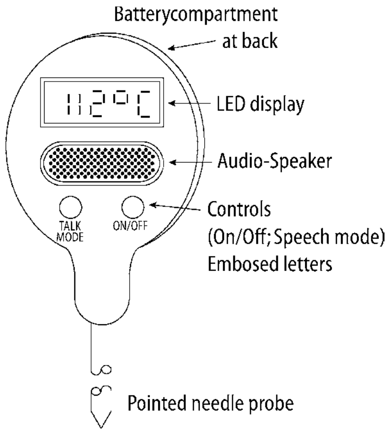


Figure 17.11. Features of a typical accessible food thermometer

Remote temperature monitoring seems to be of interest to sighted as well as visually impaired people. Unfortunately, the available products focus on monitoring the temperature of meat and are therefore not so useful to vegetarians and vegans. The grill alert talking remote meat thermometer has a probe that is inserted into the centre of the meat and a wireless transmitter that sends temperature readings to the receiver located some distance away. A voice prompt alerts the user when the meat is 'almost ready' and 'ready'. More detailed temperature information can be obtained on the LCD display. The device can be used in the oven, on a hotplate or an outdoors grill. Some typical features of an accessible food thermometer are shown in the sketch of Figure 17.11.

17.4.6 Braille Salt and Pepper Set

A Braille salt and pepper set is available from the Braille Superstore. The wooden shakers have the letters S (for salt) and P (for pepper) embossed in Braille or large print on all four sides. The shakers are made of wood with plastic stoppers and fit snugly into a heavy-duty wood base with high sides, making it difficult to accidentally knock them out.

The Buzz electronic salt and pepper mill is a chrome finish battery operated mill that can be used to grind pepper or salt to produce fine or coarse grounds. It can be used one-handed, as it is operated by pressing a button located at the top of the mill.

17.5 Environmental Control and Use of Appliances

Independent living requires blind and visually impaired people to be able to control their immediate home environment easily. This has two aspects: the accessibility and usability of the controls of a wide range of equipment for visually impaired and blind people and the provision of feedback information about the environment in an appropriate format, so that the user can make decisions about changing settings. A related issue is the accessibility and usability of equipment, such as vacuum cleaners and washing machines, that blind and visually impaired people may want to use to improve their environment. Although the controls on many appliances and other equipment are in the form of switches, push buttons and knobs operated by touch, choosing the correct option or setting on the controls generally requires vision. In addition, there are an increasing number of appliances with LCD displays and/or light emitting diode (LED) indicator lights.

Talking and/or tactile versions of a number of appliances and other devices have been developed and a selection of the available products is discussed in this section. However, many of the products are often more expensive than the non-talking versions and it may not be possible to find appropriate talking or tactile versions of all products, particularly if the blind or visually impaired person is not a speaker of English. There is also a considerably reduced choice, which could be a problem if the user has particular additional requirements. There might therefore be benefits in the development of a multilingual speech system for appliances that could be connected to a wide range of different appliances. This might require the speech system to use Bluetooth or Wifi, which would limit the appliances it could be used with to those with Bluetooth or Wifi capability.

In the meantime, many blind people use both commercially available and home-made labelling systems to make controls accessible. Some of the labelling systems that can be used by blind and visually impaired people were discussed in Section 17.2.1. Information about the status and condition of the person's immediate environment can be obtained by the use of a probe or sensor. In this section, two different types of probes with many environmental applications are described—light probes, and colour probes. The section continues with the use of sensor devices to identify temperature and barometric pressure, both of which give information on the weather, and closes with discussion about the use of some household appliances.

17.5.1 Light Probes

Light probes can be used to give blind people with little or no light perception information about their environment, so that they know how light or dark it is and whether any light is natural or artificial. Another application is to check whether any lights have been left on and the status of on/off LED and other indicator lights on appliances and machinery in the home and workplace.

There are a number of light probes from Caretec Products. Colonino detects the different nuances of colours and announces them in a clear voice, with a number of different language options available. There are three volume levels as well as the option of using earphones. The light probe function allows the user to distinguish between natural and artificial sources of light. It runs on two AAA batteries and can also be used with rechargeable batteries. It is easily portable. The LumiTest is a light probe and contrast detector. Its artificial eye senses light intensity and contrast and announces the result. It uses several acoustic signals to indicate whether it is bright or dark and whether the light is natural or artificial. A tip can be attached to the

Figure 17.12. Kentucky light probe (photograph reproduced by kind permission of Mr. Wayne Thompson(designer) and the Kentucky Office for the Blind, USA)

top of the device to control small diodes to determine whether or not electronic devices are switched on. It is portable, low cost and runs off two AAA batteries.

The Kentucky light probe is a hand-held device that produces an audio tone (see Figure 17.12). The pitch of the tone increases with the brightness of the light. Applications include detecting indicator lights on telephones, alarm panels, printers and modems, as well as other devices in the office, home or factory that use light indicators to convey information. The current design uses the Texas Instruments TSL235 light-to-frequency converter and is smaller, lighter and cheaper than other designs. It has a good sensitivity and range, producing a 1- or 2-Hz 'clicking' sound in near darkness to near ultrasonic sound (over 20 Hz) in daylight. The device is powered by two AAA batteries and current consumption is less than 2 mA, giving a very long battery life.

It has been designed to be built on a do it yourself basis. The cost of the parts is low, the circuit is very simple and the device can be built in a short time. The parts are as follows (see Figure 17.13):

- TSL235 light to frequency converter from Texas Instruments.
- AT-17 audio transducer from Projects Unlimited.
- SW412-ND pushbutton switch from Omron.

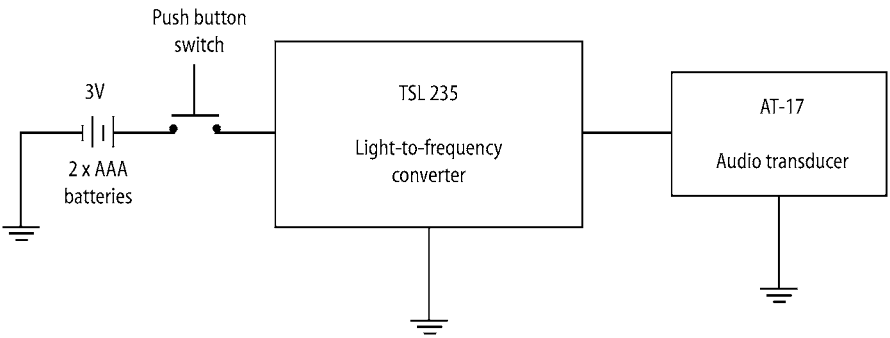


Figure 17.13. Kentucky light probe schematic diagram

- BH2AAA-PC-ND battery holder, two AAA batteries, pc mount from Memory Protection Devices.
- Two P265-ND carbon zinc AAA batteries.
- Kentucky Light Probe printed circuit board from Southland Mf.
- Small nipple PBN from Cutler's Supply.

17.5.2 Colour Probes

Colour probes can be used to identify colours. This is important both to give blind and visually impaired people information about their environment and in a number of different applications. They enable users to select their choice in colour combinations for clothes and furnishings, whether colour coordinated, contrasting or clashing, and to sort laundry by colour. They also enable users to have folders and other office materials that are coded by colour. Colour probes also have applications in education and research, for instance to inform a blind or visually impaired person of the colours of the products of a chemical reaction.

The colour of light is determined by the frequency of the electromagnetic radiation, with the colours at the violet end of the spectrum having higher frequencies than those at the red end. The basic principle of a colour identifier is the use of a sensor to detect the colour of an object by the amount of light absorbed. For instance, black absorbs all the light, whereas bright white absorbs very little light. Colour identifiers are generally able to identify a few thousand of the 2–4 million colour possibilities. This is generally more than sufficient for the colours in common use.

Accuracy depends on a number of factors, including the amount and source of lighting, including whether it is natural or artificial, in the area where the device is being used, the texture and density of the item and the quality of the technology. The precision with which the colour needs to be determined depends on the application. Some scientific applications may require very high precision, whereas determining the status of an LED may just require a distinction to be made between red and green. However, 100% accuracy may be required in making this distinction. Colour coordination applications may require both high precision and high accuracy to avoid combining shades of the same colour that do not go together. This raises issues of how the colour information is conveyed to the user, as most languages do not have a few thousand different colour words that are generally known and have a precise meaning. Expressing the colour in terms of frequency would resolve the problem of precision, but would be excessively technical and therefore off-putting as well as meaningless to most users. The different colour probes resolve this problem in different ways.

Specific colour probe products

There are a number of different colour probes, some of which also have other functions, such as light probe, calendar and clock. The number of colours that can be detected as well as how they are described also varies considerably.

The Speechmaster Colour Detector from Cobolt Systems is a commercially available colour identifier. It has both colour identifier and light probe functions. It has three volume settings and is available in several languages. Announcements can be given through the provided earphone or over the integrated loudspeaker. The device has been designed for both right and left handed use. Instructions are given in large print and audiocassette formats. A plastic snap-on cover is used to shield the detection surface when the unit is not in use. The inside of the cover is white and is initially identified before identifying other colours. The unit will repeat the colour identified continuously when switched on. This means that several repetitions of a colour have to be heard to ensure that it is identifying the current colour rather than a previous one. The unit will add 'light', 'very light', 'dark' or 'very dark' to the basic name of the colour.

The ColorTest products from Caretec are precision colour identifiers that are able to distinguish more than 1700 different colours and determine both the luminosity and hue. They have a light probe function, contrast/brightness measurement, battery statement announcement and volume control. The most elaborate of the 2000 family, the ColorTest Memo, also announces the time, date and temperature and can be used as a calendar, timer or stopwatch or to record personal memos. It is about the shape and size of a remote control. Three dots at the sensor end act to identify it and give a firm grip when the sensor is pressed against a coloured object. The battery lasts for several weeks of consistent use and the device announces clearly when it needs to be connected to the charger.

The system includes the ColorTest, a leather carrying-pouch, an AC adapter and the manual in regular print and on audio cassette. There is no earphone or wrist strap. The Braille manual is currently only available in German, though the ColorTest standard has an English Braille manual. When the sensor is held firmly against the surface to be identified and the first ('measuring') button is pressed, the colour's name is announced. Pressing the second ('analysis') button gives additional information. It can be programmed to provide varying amounts of information, including luminosity (brightness) and saturation (richness or intensity) as well as the percentages of the primary colours (red, blue and green). This additional information can be helpful in matching colours and deciding whether two colours go together. The texture and density of the surface being examined can have a significant impact on the colour being identified.

The Color Teller from Brytech is an updated version issued in 2004. It is designed solely to identify colours and has three voices for English, Spanish and French. The package includes the Color Teller, a carrying strap, a leather carrying case and documentation in both large print and audio CD in the three languages. The volume can be adjusted and the level of detail selected as simple, for example, brown or blue, or detailed, for example, dark brown or light blue. The battery installed should last for 10,000 colour announcements. The Color Teller is easy to use and reliable for obtaining a general idea of colour.

Typical features of high-tech colour indicators are shown in Figure 17.14.

There is also a simple low-tech solution to identifying colours. Colour indicator buttons from the RNIB can be used to identify items and help colour coordinate clothes. There are 16 different shapes, each of which is a different colour. Each

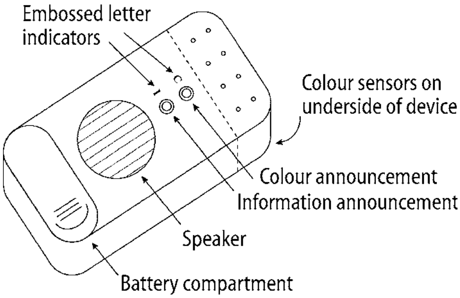


Figure 17.14. Typical features of a high-tech colour indicator

shape has two centrally punched holes. They withstand dry cleaning fluids, boiling water and the heat of an iron, but irons at the highest setting should be avoided.

17.5.3 Talking and Tactile Thermometers and Barometers

There are a number of talking thermometers, some of which have additional functions, including measuring humidity and time. The thermometer from Cobolt Systems has a probe that can be led through a window to a point outside the house. The device will announce inside and outside temperatures in either Centigrade or Fahrenheit. It can be set to do this every hour or at a selected time. An alarm can be set to alert the user to temperatures below 0 °C or above 38 °C. The maximum and minimum temperatures either daily or between set times can be stored. The RNIB distributes an inexpensive talking room thermometer key ring that tells the current temperature at the press of a button. However, it does not function at temperatures below 0 °C.

The Oregon Scientific Talking Weather Station Forecaster from Oregon Scientific provides atomic time and weather data in a single unit which can be wall mounted or placed on a desk or table top (see Figure 17.15). It provides audio broadcasts

Figure 17.15. Oregon scientific talking weather station forecaster (photograph reproduced by kind courtesy of Oregon Scientific Inc., USA)

of time, temperature, humidity and weather forecast information. The unit has a radio controlled atomic clock and five-channel outdoor temperature and humidity capability using RTGR 328 sensors or three channel capability with THGR 268 remote sensors. One RGTR remote sensor is included to enable temperatures to be measured at distances of up to 328 ft (100.0 m). The THGR 268 remote sensor has a maximum effective transmission range of 100 ft (30.5 m) line of sight. An LCD console displays weather forecasting icons, time, day and date, temperature and humidity readings and a barometric pressure bar chart. The console is updated every minute using data from the remote sensors. Hi-glo back lighting is used to make it easier to view the console at night. A pre-alarm function sounds when the remote sensor falls to 35.5 °F (0.8 °C) or below. The minimum and maximum barometric pressures, 24 h data and trend data can be displayed. The user can select whether the temperature is measured in Fahrenheit or Centigrade. The unit is powered by a 6 V AC adaptor and can be backed up by three AA batteries (which are not included). The RTGR 328 and THGR 268 sensors are each powered by two AA batteries, which are not included.

The Tissot T-Navigator watch is a multifunction device with compass, altimeter, barometer, chronograph, thermometer, alarm and countdown timer. It adapts automatically to summer and winter times until 2099 and there are two separate alarms programmable over seven days with a choice of five melodies. It uses touch screen technology and tactile sapphire glass and is water resistant to 30 m. There are various models with slight differences in the details, particularly of appearance.

The WS-2500 from La Crosse Technology is another multifunction device comprising a tactile screenclock, a barometer, a rain meter and indoor and outdoor thermometer. The clock is radio controlled by the Frankfurt antenna and there is also a manual setting function and a calendar with day and date display. It can measure indoor and outdoor temperatures between −30 °C and +70 °C with a resolution of 0.1 °C in either Fahrenheit or Centigrade. It also records the minimum and maximum temperatures.

17.5.4 Using Appliances

A number of accessible household appliances were described in the section on food preparation and consumption. However, appliances are also used for laundry, household cleaning and other household maintenance activities. Some accessible appliance are described in this section.

17.5.4.1 Washing Machines

Currently available talking washing machines seem to have been developed to help sighted people deal with complex programming cycles rather than for accessible use by blind and visually impaired people. This raises the issue of whether blind (and sighted) users would prefer to return to simpler washing machines with fewer options and raised markings on the single dial.

The first talking washing machine was developed in India and introduced in 2002. The Washy Talky from Electrolux is able to guide users through the washing process in English, Hindi and Tamil and asks the user to close the lid if it is accidentally left open. Fuzzy logic (Bandemer and Gottwald 1995; Jamshidid *et al.* 1992) is used to assess the load weight and choose the optimum programme to save water, electricity and detergent. The user is alerted if program errors occur. A digital load imbalance system is used to correct an uneven distribution of clothes. Although the system was originally aimed at families, it was found to be particularly popular with single young men, who were happy to be told how to do the washing. Electrolux is intending to obtain a worldwide patent.

The Zanussi-Electroluz Timeline ZWV 1651S Voice explains the variety of programme symbols, as well as confirming the user's actions. The machine provides 24 different functions, a 7-kg load limit and 1500 rpm spin speed.

A survey by the RNIB in 2006 found that none of the washing machines currently supplied in the UK is fully accessible to blind people, though many of them could be adapted. For instance, Hotpoint produces Braille control panels for its washing machines, dishwashing and cooking products. As far as possible, a single stud or dot is used to mark a programme or setting. These symbols are in the form of aluminium rivets, which are hard wearing and do not move easily when heated. Tactile markings can also be created using the RNIB's bump-ons, which are raised self-adhesive bumps in a range of sizes and colour. Servis have plans to introduce a washing machine with Braille markings.

17.5.4.2 Iron

Cobolt Systems distributes an electric iron guard with two polyethelene rods surrounding the edge of the iron. These act as a guide for the free hand and protect the hands from burns. The guard can be attached to any iron.

17.5.4.3 Talking Vacuum Cleaner

There are two talking vacuum cleaners in LG's Cyking Range. The two models will confirm to the user that the machine has been turned on or that the power mode has been manually altered and also tell the user when the bin is ready to be changed. The vacuum cleaners have a range of up-market features, including a high efficiency particulate filter and a silencer system.

Dyson's has developed a cheap, fast electric motor which has so much computing power that there is sufficient available to also function as a new diagnostic aid. When an appliance malfunctions, the user holds the telephone receiver up to the appliance and a speaker inside it will transmit the information required by the manufacturer, such as the serial and model number and the temperature conditions the motor has been used in. The system has now been incorporated into a prototype vacuum cleaner, which is considerably more powerful than existing products.

17.6 Money, Finance and Shopping

Being able to access money in all its forms (including cash, debit and credit cards and cheques) is very important for independence. The fact that there are a number of different currency systems in use does not make identifying different notes easier. Unfortunately, most of the existing devices work for only one set of currencies, which is not very convenient for blind people who travel abroad. There is therefore a need for a device that is able to recognise, at least, the most commonly used currencies. While many of the note identification systems perform well for new notes, accuracy is reduced for worn and damaged ones. In some currency systems, such as pounds sterling, notes of a different denomination are of a different size, making it possible to distinguish them by touch. However, this is not the case with all currency systems. For instance, US dollar notes of different values are all of the same size, though the coins can be identified by touch (Brunson 2003). Since coins are of much less value than notes, it is would seem even more important that blind people should be able to easily identify notes. Therefore, a useful solution would be the redesign of currency systems to make notes of different denominations easily recognisable by touch.

The following features could be used to enable blind and visually impaired people to distinguish banknotes (Bauer *et al.* 1995):

- Variation in size of the note with denomination, preferably of both length and height, rather than just one dimension.
- Large high contrast numerals on a uniform background. This would require a large, open space in the banknote design to give space for a numeral at least half of the banknote height. Having this numeral in a distinct area could be used to indicate banknote orientation, which is not indicated by different sized banknotes.
- A different predominant colour for each denomination of note, on at least one face of the note. This colour should be distinguishable from the other colours even in low lighting. Restricting this colour to only one side would allow the other side to be marked with a large high contrast numeral.

There may also be a role for the use of durable tactile features, printed in transparent ink (Bauer *et al.* 1995). The new notes would then have to be introduced over a period of time, as well as clearly distinguishable from the old ones to avoid problems in the transition period. There would still be issues of distinguishing between notes from different currency systems. However, it might be possible to use a combination of different types of features to do this. For instance, durable tactile marks and a large high contrast numeral could be used to determine the currency and size, predominant colour and large high contrast numerals to determine its denomination. There are both electronic and mechanical money identifiers.

17.6.1 Mechanical Money Indicators

The Right Money Gauge is a plastic device with an attached key ring to help identification of UK notes and coins. Bank notes can be matched by the height and length of the ridge, whereas coins of different denominations give different sounds.

The Pocket Money Brailler from the Braille Superstore can be use to Braille label notes with currency units of 1, 5, 10, 20, 50 and 100. Although it is produced in the US and intended for dollar notes, there is no reason it could not be used for other currencies. A corner of the note is put into the Brailler and the appropriate number pressed. The Brailler then marks the note so that it can be identified by touch. The unit is provided with an attached metal ring for clipping to a keyring or can fit into a pocket. It can label in both print and Braille.

17.6.2 Electronic Money Identifiers

The bank note money detector from Cobolt Systems indicates the value of a bank note by the number of vibrations. It is designed to enable the user to verify the value of banknotes when paying for goods or receiving change. It is very small, light and unobtrusive and available in both sterling and euro models. It is powered by a CR2032 lithium battery.

The Note Teller from BryTech is a portable bank note reader for all the old and new designs of US bank notes. Bank notes can be inserted and read in any orientation. The announcements can be provided in either English or Spanish and the volume is adjustable. The standard headphone jack can be used to provide privacy. It is battery powered with a long life battery and instructions are provided in both large print and audio CD. The Enhanced Note Tell 2 for deafblind people provides the information as a series of vibrations.

Typical features of an electronic monetary note identifier are shown in Figure 17.16.

17.6.3 Electronic Purse

An electronic purse is a chip-based application on a card that stores a monetary value in a single currency. Although electronic purses could be used to replace cash, the main applications are likely to be in special areas. They can be topped up at a bank cash machine. Mobile telephones can also be used as a form of electronic purse in some countries, with the cost charged to the telephone bill.

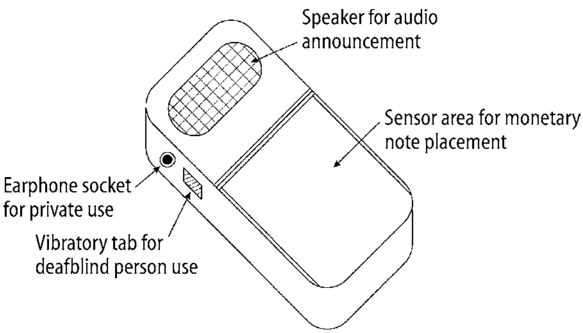


Figure 17.16. Typical features of an electronic monetary note identifier

Electronic purses were first introduced in the first half of 1990. There have been positive moves towards standardisation and the organisations providing more than 90% of currently available electronic purse codes have endorsed the common electronic purse specifications. However a number of pilot schemes have been closed due to low take-up and/or low usage, whereas those that are still running have had limited success (Hove 2005). Although electronic purse schemes may have particular benefits for blind and visually impaired people, they have generally been introduced for sighted rather than blind and visually impaired people and are only likely to be successful if there is reasonable takeup amongst the general population.

There are currently about 20 different domestic purse products available in Europe, none of which can be used internationally. Europepay is trying to promote mutual acceptance between the different schemes by the creation of minimum standards and agreements. The systems vary significantly depending on the type of organisations operating the service. For instance, financial institutions have always been particularly concerned about security, whereas public transport operators are more concerned about speed of operation. There are also issues as to whether and how individual transactions are recorded. In particular, recording them on the terminal or centrally raises privacy and security issues relating to who has access to the information and what it is used for. On the other hand, the presence of appropriate records would allow a lost card to be cancelled and the remaining value to be reimbursed when someone subsequently tries to use it. In practice, since the central system is frequently offline and updated overnight, there is generally a period of up to 24 h in which a stolen card could be used.

Recommendations by the RNIB to ensure that electronic purse systems are accessible include the following (Noonan, undated; Gill 2006):

- A standard keyboard layout, preferably the telephone layout with the number keys 1, 2, 3 at the top.
- A high contrast display with large characters and a clear typeface with easy to distinguish characters, such as Screenfont.
- A funnel opening to guide the card into the reader.
- Large buttons with clear visual markings, a single raised dot on the number 5 key and tactual feedback.
- The use of embossed symbols in accordance with the draft standard prEN 1332-5 (CEN 2006b).
- A tactile identifier of card orientation, for instance a 2-mm notch on the trailing edge in accordance with standard EN 1332-2 (CEN 1998).
- The ability to get close to the terminals so that people with low vision can read the display.
- The possibility of speech output, with the requirement coded on the user's card, in accordance with standard EN 1332-4 (CEN 2006a).
- The coding of all user requirements onto the card in accordance with standard EN 1332-4 (CEN 2006a).

- Braille is useful for some people, but not a total solution as the number of blind people who read Braille is very low. Tactual sensitivity reduces considerably with temperature and it is therefore less useful outdoors in cold climates.
- Training schemes in the use of the purses and the schemes they link into with clearly written instruction booklets in large print.
- The electronic purse system should be made as simple as possible and communications between the devices, terminals and the user should be direct and simple.

17.6.4 Automatic Teller Machines (ATMs)

An automated or automatic teller machine (ATM) is an electronic computerised telecommunications device that gives secure access to bank accounts without the need for a human bank teller or cashier. Some ATMs only have one function, such as cash dispensing, whereas others have multiple functions, such as cash dispensing, accepting deposits and printing statements (Anon 2006b). Braille overlays can be used to make ATMs accessible to blind and visually impaired people with some knowledge of Braille. Using overlays of this type does not require fluent Braille reading abilities, though users would need to be confident in distinguishing different numbers and some other symbols. The overlays are made of heavy stainless steel to prevent corrosion and applied with strong adhesives to prevent removal. They use both graphics and Braille to indicate the ATM functions. SysTech Displays produces versions for a number of different ATM models. A standardised layout and choice of functions for ATMs would also be helpful to many blind and visually impaired people with some vision, since this would enable them to know where to look for different functions. Standardisation of design between different ATMs would also make them easier to use by visually impaired people (Quarrie and Howarth, undated).

Audio lead through technology can be used to give speaking ATMs. The user plugs a set of personal headphones into a universal audio jack fitted to the front of the ATM. The automated voice guides visually impaired users to the exact locations of items such as numbers on the keypad and the cash dispenser. It also talks users through each stage of the process. Talking ATMs have audio-jacks, whereas non-talking ATMs do not. Some talking ATMs have keys marked in Braille and others have a raised dot on the number 5 key to orient users. At least some talking ATMs in the USA are identified by Braille signage and a tactile headset symbol. The technology has been available at least since the late 1990s, though there has been little evidence of its use until very recently.

The Royal Bank of Canada in Ottawa, Canada installed the world's first *talking* ATM in 1997 because of issues raised by two blind customers with the bank since 1984. However only 50 talking ATMs had been installed in Canada by the summer of 2001. In the USA, Wells Fargo agreed to install a talking ATM at all of its ATM locations in Canada in a legal settlement with blind community leaders and has subsequently installed talking ATMs at all ATM locations in all states. Bank One had installed more than 1000 talking ATMs in the US by February 2004 and Bank of America has installed a number of talking ATMs. A number of other US banks have plans to install talking ATMs or have done so. By June 2005, most of the larger banks in the UK had begun to install talking ATMs. Most recent machines are using a standardised approach based on standard PC hardware and a standard audio jack. In the USA and Canada this standardised approach is the result of regulation. The first public touch screen interface without a keypad to offer access to blind people was a talking ATM installed by Citibank in July 1999. Despite the benefits of talking ATMs, there are also concerns by some visually impaired people that the use of an audio device would indicate that the user was visually impaired and therefore an easy target for theft (Quarrie and Howarth undated).

17.7 Communications and Access to Information: Other Technologies

This section will consider the accessibility issues associated with communication and information technologies, other than telecommunications, computers and the Internet and print media and how these issues can be resolved. In particular information kiosks and smart cards are discussed and EZ Access® is presented as one potential solution to information kiosk accessibility.

17.7.1 Information Kiosks and Other Self-service Systems

Self-service systems are increasingly being used to provide a wide range of services. These include information kiosks, ATMs and machines for purchasing tickets. These systems can provide a quick easy to use service or lead to intense frustration, depending on the design. There are also issues of user preferences and the accessibility or inaccessibility of facilities to particular groups of (disabled) users.

Information kiosks based on personal computer technology can be used to give people access to a wide range of information, as long as the design is accessible and easy to use. The kiosk and all its facilities should be physically accessible, which can be difficult when trying to meet the different needs of people who are very short, those who are very tall and wheelchair users. The controls and input and output slots should be placed separately from each other, be easy to grasp, very hard to confuse and not capable of activation by mistake. The user needs to be able to plug in individual devices such as earphones, a Braille display or a wheelchair-based control device, such as an infrared link. Swipe readers should not be used in public terminals, as coordinating the swipe speed and direction and card orientation can cause difficulties for some users. Card readers may be difficult to use due to the need to identify the correct card and its correct orientation and find the slot to insert it into the reader.

A user friendly system should give the user reasonable control and the following facilities:

- A consistent method of navigation between screens and between the different levels of information and services.

- A logical and easy to follow arrangement of menus and/or options.
- The ability to easily customise the system.
- The ability to access the system using sight, hearing or touch or a combination and preferably to choose between them.
- Feedback in tactile, audio (preferably not by beeps) and visual form, as well as the ability to shut off some or all of the feedback modes.
- Being asked to confirm actions as a default, as well as the option not to be asked to confirm actions.
- Choice of the speed of operations. Many users will be discouraged from using public terminals if they feel rushed, whereas others may feel stressed if the process if too slow.

Publicly available keypads require the following properties for some degree of accessibility (Gill 1997, 2006):

- Physical buttons that can be pressed.
- Auditory and tactual confirmation of a button press (with the option to turn off the auditory confirmation if the transaction is likely to last any time).
- Telephone style layout, with the digits 1, 2, 3 on the top row of the number pad.
- A dot on the number 5 key.
- Tactual differentiation and clear separation of the number keys and the function keys.The command keys should be to the right of the numeric keys, as keys in this position are less likely to be inadvertently pressed.
- Characters should be of a reasonable size and have good contrast with the colour of the key *e.g.* black characters on a matt white background.

Screens should have the following properties to increase accessibility (Gill 1997, 2006):

- A large screen with clear graphics.
- Protection from glare.
- Viewable from a wheelchair.
- A legible typeface with text on a plain background and the possibility of increasing the text size.
- Avoidance of scrolling or flashing text.
- The use of standard icons, as far as possible.
- No barriers to users getting close to the screen.
- Good colour contrast with the background and avoidance of colour usage which could cause problems to colour blind people.
- An adjustable display angle.
- The option of selecting the language. Preferably all the languages spoken by significant numbers of people in the area as well as the main languages likely to

be spoken by tourists should be included. There are then issues of how a possibly large number of choices should be presented.

Examples of accessible information systems

TouchVision in California has developed a talking touch-screen kiosk using text-to-speech technology to read the screen content. The content and available choices are the same whether or not users use the speech-enhanced option.

A way-finding talking kiosk in New York runs on an internal computer and combines a tactile/talking map with a standard telephone style keypad. This gives travellers a choice of options for accessing way-finding information which is delivered as both speech and output on a high contrast large print video display. The kiosk alerts potential users through bird song which can be heard from 150 ft (45.7 m) away. The user is then greeted when they come within 10 ft (3.0 m) of the system and welcomed and introduced to the system from 2 ft (0.6 m) onwards. Users who are familiar with the system can hit the star key on the keypad to go to the main menu. The system is being used by both sighted and visually impaired people.

Some terminals in Woolworths Supermarkets in Australia have been made accessible by putting a Brailled overlay over the keypad and a tactual representation on the under-side of the machine to tactually inform user of the correct card orientation.

The information kiosk developed to provide information on the museums in Tampere, Finland is placed on a stand with a web camera on top of the kiosk. A touch screen is used to navigate the kiosk. The main area of the user interface is used to provide the information requested. There is an interactive agent in the right upper corner of the interface and a video image taken by the web camera below it. The agent is a talking head which shows facial expressions and can turn and look in different direction. Different buttons can be used to give commands to the kiosk. The position of the components can easily be changed by editing the configuration file. Information is provided in Finnish, but the language could easily be changed. The talking head has computer vision, allowing it to detect potential users and invite them in and greet and say farewell to users as they enter and leave. Some users will like this human type touch, whereas others may find it annoying. The system still requires further development, as the talking head sometimes loses visual contact with users and then greets them again when detecting them. The repeated greetings can become very irritating (Mäkinen *et al.*, 2002).

17.7.2 Using Smart Cards

A smart card is a credit card sized plastic card that incorporates an integrated electronic circuit. There are four main types:

- Memory only: they are often used as prepayment cards for public telephones.
- Microprocessor: they can include security features, for instance for financial applications.

- Contactless: the card does not need to be put into a reader and can work from a distance of up to 10 cm. This is useful for transport applications. Contactless cards also have a reduced risk of the card being stolen during use.
- Combined contact/contactless: they are more expensive to make, but are useful in some multi-application systems.

Smart cards should have embossed symbols to enable blind people to select the correct card and there should be a 2-mm notch on the trailing edge to make it easier to insert the card correctly. Biometric systems use physiological or behavioural characteristics to identify the user rather than requiring them to enter a number. This will make the system easier to use for some people, but may raise security and privacy issues, particularly when the biometric information is stored centrally rather than on the smart card. There should always be the choice of an alternative verification system, such as PIN, with the decision of which system to use left to the user.

Facial recognition is probably the simplest and least intrusive system. Although not the most accurate, it gives reasonable results when used in combination with a secure token such as a smart card. Fingerprint systems are fairly accurate, but unsuitable for users with prosthetic hands or injured fingers. They may also be unacceptable to many users due to their association with the detection of crime. Iris recognition is accurate, but more intrusive. It requires users to position their eye relative to a camera which can be difficult for tall or short people or wheelchair users, as well as blind people. It may also be unacceptable to some ethnic and religious groups.

Smart cards could contain information about the user's preferred method of using a machine or terminal. For instance a visually impaired person could have a card that is encoded to instruct the terminal to display large characters on the screen, to provide audio output (preferably with a headphone to ensure privacy) and to give the user additional time or simplify the choices. The draft European standard EN1332-4 (CEN 2006a) contains a method of encoding this information. Smart cards are relatively widely used and therefore do not have the disadvantage of marking users out as obviously 'disabled'. However, the use of any device, such as smart cards, which encodes information about the user raises important privacy and security issues, particularly as the user cannot necessarily access this information or determine if they have accessed all the information about themselves.

A Special Needs Application Program Interface (SNAPI) can be used to put a user's preferences on a smart card and the cards can be used in cash dispensers, ticket machines and public access computers. The machine automatically returns to the default settings when the card is removed. The Snapi software standardises the coding of individual preferences to enable them to be stored on a smart card. The information stored is based on the person's preferences and is chosen by them. However, a number of machines will not have the facilities to comply. For instance most ATMs do not currently have speech capability. The person's identity should not be visible to the system, only their preferences for using it. The Department of Transport has incorporated the new standard into its specifications for ticket machines and banking organisations have said that they will implement Snapi, but have not yet indicated when this will occur.

17.7.3 EZ Access®

Touchscreens are being used much more widely in commercial and social applications. Items, such as instructions, text, pictures, films and buttons can be selected on the screen and the associated touchscreen area pressed by the user. However, the use of a touchscreen requires vision. A set of interface techniques, called EZ Access, have been developed to give disabled people access to touchscreens. It consists of software and hardware enhancements that provide an adaptable touchscreen interface and allows users to choose the modalities that are used to present information. The touchscreen strategies used with EZ access can be extended to other electronic devices. An infrared link can be used to provide indirect access. A number of public information kiosks have EZ Access (see Figure 17.17).

EZ access operation modes include the following (Law and Vanderheiden 1998):

- List mode: all the functions (buttons, displays, controls) are available in a list that can be navigated easily using the three commands to move up the list, to move down the list and to select an item.
- Select and confirm mode: a button or control is selected and announced, but not activated when it is touched. In order to activate it a confirmatory action, such as pressing another button or holding the button down, is required. This allows the user to explore the names and functions of buttons without activating them until the desired one is reached. Help text is sometimes read out when the button is held down or pressed repeatedly.
- Auto-step scanning mode: the items are highlighted or announced sequentially until the user indicates that the desired item has been announced and this item is then activated.
- Showsounds mode: all audio information is also presented visually.
- Direct text control techniques: users can navigate and activate controls using text only.
- Auxiliary interface port: users can connect personal devices and use them instead of the controls or displays on the device.
- Standard mode: the mode that the device would use as a default.

Blind users can use an audio list mode, with all items on the screen organised in a list which is displayed vertically at the edge of the screen. Listed items are spoken aloud as the user slides their finger down the screen and they press a confirmation button, generally located below the screen, when they find the desired item. Users with low vision could use an audio select and confirm mode or an audio list mode. When touching objects, they are highlighted and spoken aloud. Deafblind users could attach a Braille notetaker or other tactile device to the infrared port and use it to access the kiosk.

Figure 17.17. Postal service EZ access self-service kiosk in USA (photograph reproduced by kind permission of University of Wisconsin-Madison Engineering External Relations, USA)

Quad Media has developed a voting kiosk with EZ Access. It runs on Windows and enables users to view or receive information using different modalities, including sight and hearing. The user can operate, control and interact with the program in different ways. The system is intended to be easy to learn with cue operations that help the user and does not require fine motor control or the use of vision. The kiosk has text-to-speech capability and is able to read aloud written questions.

17.8 Chapter Summary

This chapter has used the activities module of the comprehensive assistive technology model to classify the accessibility devices and solutions available to overcome the barriers encountered by visually impaired and blind people in carrying out *daily living* activities. As well as providing a useful structure for the chapter, this approach has provided an illustration of the value of the CAT model in discussing and describing assistive technology provision. Therefore, the main sections of the chapter are the five daily living activities in the CAT model: personal care, time-keeping, alarms and alerting, food preparation and consumption, environmental control and the use of appliances and, finally, money, finance, and shopping.

In personal care, the two areas identified as posing accessibility barriers for visually impaired and blind people were those of identifying personal items such as clothing, and being able to monitor their health status and use healthcare products independently. Currently available solutions comprise advanced and low technology labelling systems and a range of accessible healthcare devices and systems. Six types of accessible healthcare devices were described in the chapter.

Time keeping, alarms and alerting is the second category of daily living activities in the CAT model. This section described a range of clocks and watches with tactile or audio interfaces for telling the time. Some of these devices are multifunctional and incorporate calendars or other features. Alarms and alerting devices are an important feature of the modern home environment and three alarm systems and one locator device were described in this section.

The food preparation and consumption section of the chapter focussed on assistive technology to make food preparation accessible. The section comprised descriptions of a range of high and low-tech devices, including talking kitchen scales, liquid level indicators and a Braille salt and pepper set.

The fifth section of the chapter discussed environmental control and the use of household appliances. A number of sophisticated high technology solutions in the form of light and colour probes are available for determining the environmental parameters of light and colour. Applications of colour probes range from identifying the colour of personal clothing to determining whether a chemical reaction has occurred through a change of colour. In the former case, there are also a number of low technology solutions. Accessible appliances for use in household laundry and cleaning were described in the latter part of this section.

The high- and low-tech solutions to making money, finance, and shopping accessible to blind and visually impaired people were considered in Section 17.6. These include talking ATMs and mechanical and electronic money identifiers, as well as proposals for making different denominations of banknotes easier to distinguish from each other.

As discussed in the learning objectives, technologies for making communication and information using technologies other than computers and the Internet, telecommunications and print media accessible to blind and visually impaired people are considered in this chapter, since their applications are mainly in daily living activities. Thus Section 17.7 considered the requirements for making information kiosks and other self-service systems, as well as smart cards, accessible to blind and visually impaired people. EZ access is presented as an example of a solution to kiosk accessibility.

Although a number of high technology solutions are available in some areas, the use of simple, sometimes homemade tools and devices is more common. This gives rise to questions about how best existing gaps can filled and in particular, the balance between the following:

- Assistive technology and *design for all*.
- The use of technology and human assistance.
- High and low technology solutions.

Other important questions relate to ensuring that solutions for blind and visually impaired people are no more expensive than the devices and products available for sighted people, are easy to obtain information about, easy and intuitive to use and come with good documentation in local languages. The language issue is very important, since solutions that are available in non-European languages are relatively rare and some devices are available only in English.

Questions

Q.1 Describe how the comprehensive assistive technology model can be used to provide a framework for categorising daily living activities.

Q.2 List the main uses of a labelling system. Describe two different technologies that are used in such systems.

Q.3 List and briefly describe the main interface technologies that are used in time-keeping devices for:
(a) Visually impaired people
(b) Blind people.

Q.4 Describe a set of accessible food preparation devices that could be used by a blind person. Identify the accessibility interface(s) for each device.

Q.5 List and briefly discuss the main areas where assistive technology is required to remove accessibility barriers to accessing and using money.

Projects

P.1 Consider the design of a kitchen which is accessible, easily usable and safe for blind people. Draw up guidelines for accessibility, usability and safety which cover the following areas:
(a) Layout
(b) Decor
(c) Equipment
(d) Cleaning
(e) Utility services

P.2 How does a visually impaired person or a blind person control the environment in their own home? Consider all the different appliances that they might want to control. Discuss the current state of accessible technology in this area. Evaluate the usability of this technology. Identify any aspects of environmental control where a blind or visually impaired person would encounter barriers with the current state of technology. Draw up detailed proposals, including performance and end-user specifications, for technologies to overcome these barriers.

P.3 Discuss the role of high-tech assistive technology, low-tech assistive technology, *design for all* and human assistance in making different categories of daily living activities accessible to blind and visually impaired people. Use the activity component of the CAT model to chart how the different approaches could be used for different activities.

P.4 A blind or visually impaired couple are about to become parents. Consider all the activities involved in caring for small children. Identify activities where the couple might experience accessibility barriers. Evaluate the technology that is currently available to overcome these barriers. Identify areas where additional technology will be required and draw up detailed proposals, including performance and end-user specifications, for technologies to overcome these barriers.

References

Anon, 2006a, Radio frequency identification (RFID), http://en/wikipedia.org/wiki/RFID

Anon, 2006b, Automated teller machine, Wikipedia, http://en.widipedia.org/wiki/Automatic_teller_machine

Bandemer, H., and Gottwald, S., 1995, Fuzzy sets, fuzzy logic and fuzzy methods with applications, John Wiley and Sons Ltd., Chichester, UK

Bauer, N.S., Baum, G.A., Brabyn, J., Gaynor, J., Haslop, J., Legge, G.E., Shannon, R.R., Sincerbox, G.T., Church, S., Hyland S., and Prisco, J.M., 1995, Currency features for visually impaired people, National Academy of Sciences

Brunson, M., 2003, Show me the money: making currency accessible to people who are blind, http://www.rit.edu/~easi/itd/itdv09n1/brunson.htm

CEN, 1998, Identity cards, cards, interfaces, ergonomics, identification methods, dimensions, orientation, position, edge, indentations, shape, EN 1332-2, European Committee for Standardisation

CEN, 2006a, Identification card systems – man-machine interface – part 4: coding of user requirements for people with special needs, Draft EN 1332-4, European Committee for Standardisation

CEN, 2006b, Identification card systems – man-machine interface – part 5: Raised tactile symbols for differentiation of application on ID-1 cards, Draft prEN 1332-5, European Committee for Standardisation

EC, 2005, Guidance concerning the Braille requirements for labelling and the package leaflet, Article 56(a) of Council Directive 2001/83/EC, as amended, European Commission

Gill, J., 1997, Access prohibited? Information for designers of public access terminals. http://www.tiresias.org/pats/browse.html

Gill, J., 2006, Guidelines for the design of accessible communication and information technology systems, http://www.tiresias.org/guidelines/index.htm

Hove, L.V., 2005, Electronic purses (which) way to go?, First Monday Special Issue 3: Internet banking, e-money, and Internet gift economies, http://www.firstmonday.org/issues/issue5_7/hove/.

Jamshidi, M., Vadiee, N., and Ross, T.J., 1992, Fuzzy Logic and Control, Prentice-Hall

Law, C., and Vanderheiden, G., 1998, EZ access strategies for cross-disability access to kiosks, telephones and CVRS, CSUN, http://www.dinf.ne.jp/doc/english/Us_Eu_conf/csun_98/csun98-07...

Mäkinen, E., Patomäki, S., and Raisamo, R., 2002, Experiences on a multimodal information kiosk with an interactive agent, NordiCHI

Noonan, T., undated. Supplemental HREOC e-commerce reference response, Blind Citizens Australia, http://www.hreoc.gov.au/disability_rights/inquiries/ecom_subs/bca1.htm

Quarrie, C.J., and Howarth, P.A., undated, The use of smart cards by the visually impaired. http://www.lboro.ac.uk/departments/hu/groups/viserg/publications/smart.htm

Resources

Engelen, J., Fletcher, R., Gill, J., Klusell, L., Linström, J-I., Silver, J., and Sutherland, B., A vision of financial services, access to financial services for visually disabled people, RNIB

Gill, J., and Devine Wright, H., 1999, Selecting cards by touch, RNIB, http://www.tiresias.org/tdiff

Organisations that provide accessible daily living products and devices for blind and visually impaired people

ABLEDATA, product database, USA, http://www.abledata.com
Adaptive Technology Consulting, USA, http://www.adaptivetech.net/products
American Printing House for the Blind, USA, http://www.aph.org
American Thermoform Corporation (ATC), USA, http://www.americanthermoform.com Braille supplies and equipment
ASSISTec Ltd, Israel, http://www.assis-tec.com
Braille Superstore, Canada, http://www.braillebookstore.com Braille and talking products
Brytech, Canada, http://www.brytech.com/ Banknote readers and health monitoring products
CareTec Gmbh, Austria, http://www.caretec.at
Cobolt Systems Ltd, UK, http://www.cobolt.co.uk
Diabetes Action Network of the National Federation of the Blind, USA, http://www.nfb.org/voice.htm
Dynamic-Living.com, http://www.dynamic-living.com
En-Vision America Inc, USA, http://www.envisionamerica.com/ Technology for people with visual and cognitive impairments
Exceptional Teaching Aids Inc, USA, http://www.exceptionalteaching.com
Eye-Dea Shop, USA, http://www.clevelandsightcenter.org/shop.csc/
Ferguson Enterprises, USA, http://www.fergusonenterprises.com
Homecraft AbilityOne, UK, http://www.homecraftabilityone.com
Independent Living Aids Inc, USA, http://www.independentliving.com
Institut de Réadaption en Déficience Physique de Québec, Canada, http://www.irdpq.qc.ca/
Kentucky Department for the Blind, USA, http://blind.ky.gov/
KGS Corporation, Japan, www.kgs-jpn.co.jp
La Crosse Technology, Australia, http://www.lacrossetechnology.com.au
Lighthouse International, USA, http://www.lighthouse.org
MAB Community Services, USA, http://www.mablind.org
National Association for the Visually Handicapped, USA, http://www.navh.org
National Federation for the Blind, USA, http://www.nfb.org
Oregon Scientific, UK, http://www.oregon.scientific.co.uk
Quad Media, USA, http://www.quadmedia.com/
Royal National Institute for the Blind (RNIB) (online shop), http://onlineshop.rnib.org.uk/
RVB Systems Group, USA, http://www.barcode-solutions.com/
San Francisco Network of Support for Community Living, USA, http://sanfrancisco.networkofcare.org
Spektra v.d.n., Czech Republic, http://www.spektravox.cz
SysTech Displays, USA, http://www.systechdisplays.com
Ultmost Technology Corporation, Republic of China, http://www.asia.globalsources.com/ultmost.co

18 Assistive Technology for Education, Employment and Recreation

Learning Objectives

The comprehensive assistive technology (CAT) model introduced in Chapter 1 has three fundamental activity categories and three contextual activity categories. The assistive technologies required to support blind and visually impaired people in carrying out activities in the fundamental activity categories were discussed in Chapters 5–16 and the assistive technologies used in the contextual activity category of daily living were presented in Chapter 17. This chapter will consider the assistive technologies required for the remaining two contextual activity categories of education and employment and recreational activities. The decomposition of these activities in the CAT model is used to provide the structure for this chapter. As already indicated in Chapter 1, many of the assistive technologies required to support blind and visually impaired people in education and employment are those required to access print media, computers and the Internet and telecommunications and have therefore been discussed in earlier chapters. However, the discussion in this chapter will show that there has not been a systematic effort to develop assistive technologies to overcome accessibility barriers in education, employment and recreational activities. Therefore there are numerous gaps in provision and potential users are often not aware of what is available. The learning objectives for the chapter include the following:

- Using the CAT model as a framework for categorizing the assistive technology used in education, employment and recreational activities.
- Obtaining an overview of the role of assistive technology in these areas.
- Understanding the engineering principles of some of the assistive technology solutions.

18.1 Introduction

The ability to take part in education and training, employment and recreational activities should be considered a fundamental right. Wider participation and social involvement by all sections of society is also of benefit to both individuals and society as a whole. However, the statistical data (see Sections 18.2 and 18.3)

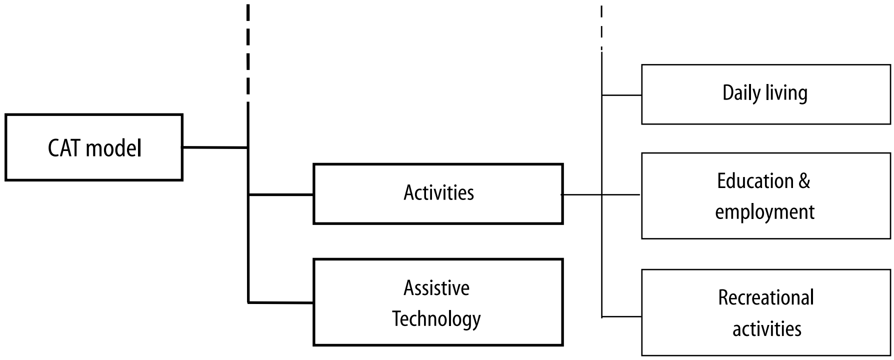


Figure 18.1. Contextual activity categories in the CAT model

indicates lower participation by blind and visually impaired people in education and considerably lower rates of obtaining qualification and employment than the population as a whole, indicating that there are significant accessibility barriers in these areas.

The *activities* component of the comprehensive assistive technology model presented in Chapter 1 is used to provide the structure for a systematic presentation of the different assistive technology solutions considered in this chapter. Figure 18.1 shows the components of this model that form the basis of this structure.

The focus of this chapter is the assistive technologies required specifically for education, employment and recreational activities. However, unless all these activities take place from home, blind and visually impaired people will require accessible transport and buildings, as well as orientation and mobility training and technology. Assistive technology for accessing print media, computers and the Internet and, to a lesser extent, telecommunications, are also generally of great importance.

Education and employment involve a wide range of different activities associated with education and training and the workplace. The full categorisation of the relevant activities is taken from the third level of the CAT model as shown in Figure 18.2. This classification is used in Sections 18.2 and 18.3 to detail the assistive technology solutions available to overcome barriers in education and employment.

Author Note. It is in the nature of assistive technology for education, employment and recreational activities that the authors will describe a number of products. These have been chosen to illustrate typical features that can be found in available products. The selection made is not comprehensive, merely illustrative. Readers seeking comprehensive lists of products are urged to consult a database like that of the ABLEDATA website or visit the websites of the many assistive technology distributors. A list of useful websites can be found in the Resources section at the end of the chapter.

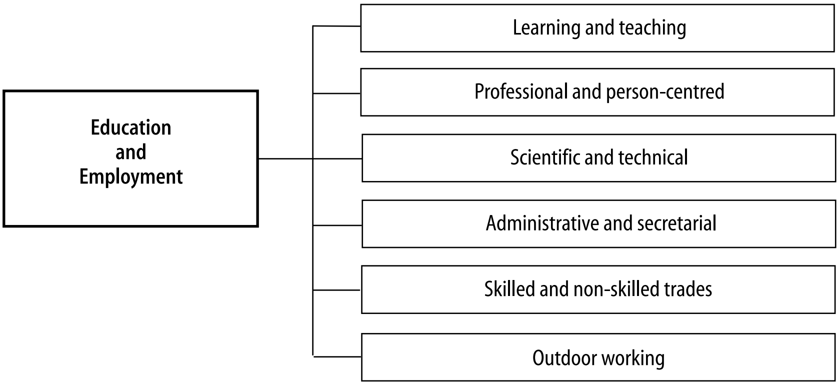


Figure 18.2. Education and employment: the third level areas

18.2 Education: Learning and Teaching

Education and training are very important for a number of reasons including their value in themselves, self-development and obtaining qualifications which are increasingly a prerequisite for any type of employment and essential for better paid jobs with a career structure. This means that both educational materials and the teaching process need to be made available to disabled students and trainees, including blind and visually impaired ones. However, blind and visually impaired people still experience barriers in accessing education, particularly at the higher levels, and acquire fewer qualifications than do sighted people. For instance, in the USA in 1994/1995, only 45% of severely visually impaired or blind people obtained a high school diploma, compared to 80% of sighted people (Anon, undated).

In the UK in 2004/2005, some 0.12% of postgraduate students living in the UK and 0.16% of undergraduate students living in the UK were blind or visually impaired. These figures have increased from 0.06% and 0.13% for undergraduate and postgraduate students respectively for 1994/1995. This percentage is considerably less than the percentage of blind and visually impaired people in the population as a whole, estimated at between 0.8% and 2.0% for people in the 16–64 age range (Tate *et al.* 2005) and 0.24% in the 0–16 age range (Keil and Clunies-Ross 2003). Although some of the difference is due to a combination of different definitions and the fact that all prevalence of visual impairment increases with age, with the 40–64 age range not well represented in the student population, blind and visually impaired people are still significantly underrepresented in higher education.

In many countries there are moves away from separate education for disabled children and young people, including for blind and visually impaired people, and towards integration in mainstream schools. This is generally to be welcomed, in terms of being a better preparation for employment in a largely sighted world and generally being able to provide a wider choice of subjects, as well as accustoming sighted people to interacting with blind and visually impaired people and contributing towards overcoming prejudices and negative stereotypes. However, it is important that educational provision is adequately resourced, including the provision of suitable and good quality assistive technology. Issues relating to the accessibility of text materials and music notation were discussed in Chapters 15 and 16, respectively and the assistive technology for using computers and the Internet was considered in Chapters 12 and 13.

In the educational context, the focus is generally on blind and visually impaired, or other disabled people as students, rather than as teachers and professors. In the teaching context, the ability to read student's work in order to comment on it and mark it, is very important. At the further and higher education level, submitting electronic versions of work, for instance in the form of a word or open office document, is often possible and would require a screen reader or magnifier to be used. However, screen readers may have problems with mathematics or scientific formulae and this does not resolve the problem of assessing hand written student work. In this case, either an electronic or a human reader will be required. The Talking Camera or Kurzweil-NFB Reader is a recently developed device that could provide a solution. It consists of a digital camera packaged with a pocket computer. It is designed to read documents and the camera has a field of view report feature that can be used to align the document and reader. Further experience with the device will be required to see how it copes with material that is difficult to read. It is also unlikely to resolve the problem of reading mathematics or scientific formulae, at least in the short term. However, it may be possible to develop an expert system or other rule-based approach to do this. The READ IT project described in Chapter 12 of this book is another potential route to solving this text accessibility problem.

This part of the chapter is divided into two main subsections on accessing educational processes and approaches, and educational technologies, devices and tools.

18.2.1 Accessing Educational Processes and Approaches

This section considers access to specific educational processes and approaches where blind and visually impaired students may encounter barriers, including graphics, mathematics, computer aided learning and e-learning.

18.2.1.1 Accessing Graphics

Graphics are used quite extensively in teaching materials, which can cause problems in making the material accessible to blind and visually impaired people. The issue here is conveying the educational content, including the content incorporated in graphics, in a way that is comprehensible to and facilitates learning by blind and visually impaired people rather than making the graphics in isolation accessible. This means consideration of the material as a whole and gives a number of different approaches, including the following:

- The use of tactile diagrams, possibly supplemented by text. Since tactile or haptic perception does not give the same overview as visual perception, the production of tactile diagrams requires the information to be edited rather than simply the production of a tactile version of the visual graphic. In particular, the amount of information should be reduced to its main features and orientation, and navigation through the graphic needs to be simplified and facilitated. In some cases, successive representations of increasing complexity and detail may be helpful. It is sometimes useful to supplement tactile diagrams by text descriptions, for instance to give an overview of the diagram and point out its main features.

- A detailed textual description of the significant features of the graphic.
- Editing the textual material, so that it contains all the educational content contained in the graphic.
- The use of a touchpad system to allow blind and visually impaired people to tactilely investigate graphics on paper and on screen. However, as discussed above, the unedited version of the graphics may be difficult to navigate and make sense of tactilely.
- The use of formalised natural language approaches to provide descriptions of data graphs (Ferres *et al.* 2006).
- The use of musical systems to represent data graphs, but this approach does not work satisfactorily when the graph is not time dependent, as in the case of pie charts and when the graph contains more than three lines. In addition, important information, such as the labels on the axes, is not presented in a purely auditory graph (Ferres *et al.* 2006).

18.2.1.1.1 Producing Tactile Graphics

There are a number of different methods for producing tactile graphics (RNIB 2006). The appropriate method will depend on the type of graphics and the preferences of the potential user(s). Commercial processes for producing tactile graphics include screen-printing, using a screen or fine mesh with a negative image. To produce a tactile graphic, resin rather than ink is forced through the screen onto the desired surface. Some signage companies that produce Braille signs also produce tactile graphics, generally in moulded or shaped plastics. Using a commercial firm is likely to result in a tactile graphic that is durable, gives more information than other methods, involves both touch and colour, and is easily mass-produced. However, this is likely to be expensive and time-consuming. Embossed graphics comprise images consisting of dots, generally produced on a Braille embosser. Although standard embossers can produce graphics, variable dot or specialist embossers generally give better results. Braille translation software such as Duxbury can import special image files to emboss them as part of a standard document and the Tiger Advantage embossers from Viewplus can emboss directly from Windows applications. The advantages of this approach are ease of production and low cost. The disadvantages are the possible poorer discrimination quality of the graphic due to the lack of height variation in the tactile images produced by most embossers and limited shapes and fills available.

Tactile-audio graphics are produced by placing a tactile graphic on top of a touch sensitive pad connected to a computer. The associated software package enables the user to define regions of an image and associate text or recorded descriptions with them. When one of the defined areas is touched subsequently, the user hears the text or recorded description. This has the advantage of giving access to additional information, but is not easily portable and requires special software and a touch-pad. Swell paper has a coating of heat reactive chemicals and has embedded microcapsules of alcohol which burst when exposed to heat, making the paper surface swell. Marking the paper with black ink before heat is applied controls the areas that are raised. Images can be transferred to swell paper using photocopiers or printed directly. Laser printers are likely to give a poor quality image due to background swelling from the heat of the printer and some inkjet printers work better than others. Alternatively a marker pen or a heat pen can be used. Various modelling techniques can be used, either professionally or by individuals. Carefully designed models can provide information on the scale and relationship between objects that is difficult to convey otherwise. A range of materials can be used, including SmellyVision from Zychem Ltd and Tactile Colour from Tactile Colour Ltd. SmellyVision comprises coloured paper sheets with scratch and sniff scents, with each colour having an associated scent. Tactile Colour comprises coloured sticky back sheets with a particular texture associated with each colour. SmellyVision and Tactile Colour are particularly suitable for graphics for children.

18.2.1.2 Access to Mathematics

Other important areas are the representation of mathematics and scientific, particularly chemical, formulae, where blind and visually impaired people need to be able to do the following:

- Read textbooks and/or the scientific literature.
- Communicate with blind and sighted mathematicians and scientists and/or blind and sighted fellow students, including by producing material containing mathematical and/or scientific formulae.
- Navigate, analyse, manipulate and calculate or otherwise modify formulae.

In addition, there are considerable advantages in blind and visually impaired people being able to use the same representation throughout their education and employment. Unfortunately, a standardised notation for the representation of mathematics for blind and visually impaired people has not yet been developed. It should also be noted that, although the representations used by sighted people are visual, mathematical concepts are not inherently visual. Approaches to obtaining mathematical representations for blind people have to date been based on the sighted notation and frequently resulted in solutions that are relatively difficult to navigate, analyse and use in calculations. However, the development of a mathematical representation for blind people that is not linked to sighted notation would have the serious disadvantage of not facilitating communication between blind and sighted mathematicians.

Braille is an obvious candidate for use in mathematical representations, but most countries have their own national versions and there is no universally accepted standard. For instance, UK Mathematics Braille is used in the UK, the Nemeth

Code in North America and the Marburg code in Germany. Although all these codes are based on six dot Braille, they have significant differences, particularly with regards to number representation, which affects the representation of other symbols (Barker 2006). There are eight dot mathematics codes, including DotsPlus Expert, in which an additional dot in the dot-7 (lower left) position is used for capitalisation and Greek letters are represented as single 8-dot cells with an additional dot in the dot-8 (lower right) position. DotsPlus Braille was developed to allow mathematics to be represented tactilely in the two-dimensional form used by sighted people rather than linearly. The MathType editor in MS Word and version 2.2 of the MathPlayer plug-in for Internet Explorer allow equations to be formatted as DotsPlus (Gardner *et al.* 2006).

Translation from visual text to Braille involves both content and formatting and it is generally easier to produce correct content than good formatting. Duxbury Systems produces the only commercially available software that translates to standard mathematics Braille. The Duxbury Braille translator (DBT) includes Math/Science Code and Computer Braille translations for US, UK and French Braille. MegaDots 2.3 includes MegaMath, which also supports chemistry symbols. MegaDots 2.3 and DBT 10.6 can both emboss to USB devices, network embossers and the ViewPlus Tiger Series (see below). However, there are a number of projects to develop mathematics Braille translators, including the following:

- Libloius, which aims to translate from MathML to Nemeth and UK Maths Braille initially, followed by various European mathematics Braille codes (Gardner *et al.* 2006).
- The Universal Maths Conversion library, which aims to create a library to support two-way conversion between mathematical notation and mathematical Braille, including Nemeth, Marburg, French (two versions), UK and Italian (Archambault *et al.* 2004).
- LaBraDoor which translates LATEX to Marburg Braille (Batusic *et al.* 1998).
- Math2braille, which is an open source architecture for translation from MathML to Braille (Crombie *et al.* 2004).
- Insight, which translates from Nemeth Braille to LATEX (Annamalai *et al.* 2003).
- MMBT (Multi-language Mathematical Braille Translator), which is an ongoing opensource project about translation from French (two versions), UK and Italian Braille to LATEX and MathML (Moço and Archambault 2003).
- SBT (Spanish Braille Translator), which aims to translate from Spanish Braille to MathML (Alonso *et al.* 2006).

The Canonical MathML (CML, undated) structure has been developed to allow each mathematical structure to be represented in a unique way and facilitate translation into Braille (Archambault and Moço 2006). It provides a stylesheet to allow conversion to Braille by implementing the rules of the target Braille code. Versions for French (two notations) and Italian are available and Marburg and UK versions are being developed (Archambault and Moço 2006).

The LAMBDA project (LBA, undated) aims to provide an editor for writing mathematical expressions in a linear and effective way that allows active manipulation, reading and computation using Braille displays or synthetic speech. The Lambda Math Code is a linear mathematical code that has been directly derived from MathML. A conversion system is provided through MathML to Latex, MathType and Mathematica and a translation program to Braille which takes account of national differences. It has developed a common markup language which can be used to represent mathematical expressions linearly in a way that preserves the features of the national Braille representations. The representation is intended to be compact to minimise finger movements over the Braille display, to preserve national configurations for the most common symbols and to use intuitive rules and dot combinations (Schweikhardt *et al.* 2006).

The Lambda software editor allows users to write and manipulate formulae. Commonly used operations, such as opening and saving files, deleting and copying, use standard Windows commands with which the user is already likely to be familiar. The representation is based on a block structure, with different blocks (for instance representing a fraction) indicated by open and close structures. The system is able to recognise blocks and provide commands for operating on an entire block, as well as to hide block content in order to highlight the structure of the expression and facilitate understanding it. This could resolve an important problem encountered by blind people when reading mathematics of trying to learn and understand the overall structure of an expression, for instance how many fractions it contains or whether the square root is being taken of the whole expression. To reduce the likelihood of structural errors in data entry related to blocks, a single command can be used to delete block markers and a generic command to close a block, with the program inserting the correct marker based on the context.

The Tiger series of embossers can be used to produce both tactile graphics and Braille mathematics with graphics. Tiger embossers can also emboss Braille directly from standard Windows software, such as MS Office, in addition to using Braille-making software, such as Duxbury DBT 10.6 and Megadots 2.3. Braille and graphics can be embossed together on standard paper and do not require thermal paper. Tiger embossers are also claimed to be relatively quiet. They are desktop size.

Other approaches include spoken mathematics. This raises issues of how punctuation and phrasing, including the grouping of particular mathematical elements can be used to facilitate understanding (Fitzpatrick 2006). WinTriangle is a specialised word processor which is able to display and speak the main symbols used in mathematical and scientific expression. Its menus and hot keys allow access to and voicing of several fonts, including the Triangle.ttf font. This font's mark-up symbols allow most mathematical and scientific expressions to be expressed in a linear form. The system supports both the speech application programming interfaces SAPI 4.0 and 5.1, giving users a choice of different text to voice engines. When used with SAPI 5.1 different voice settings can be used for standard text and mathematical symbols.

18.2.1.3 Computer Aided Learning and e-Learning

The use of computer aided learning systems and, in particular e-learning, are becoming increasingly important. This raises a number of issues, including the following:

- All e-learning materials and sites need to be fully accessible to blind and visually impaired, as well as other disabled people. Issues relating to the accessibility of web sites and documents produced in PDF format were considered in Chapter 12.
- e-Learning needs to be adequately resourced. It is often popular with governments, as it is seen as an inexpensive option. However, education provided without adequate resources in terms of staff, equipment, infrastructure and teaching materials, is generally of poor quality and is likely to disadvantage particular groups of students, such as blind and visually impaired students, with a need for specific types of resources. This is particularly important, as a design for all approach, based on the automatic design of teaching materials for the full diversity of the student population, has not yet been implemented and therefore the formats required by blind and visually impaired people are often seen as additional, including in terms of costs.
- e-Learning should not be seen as a substitute for making traditional face-to-face learning accessible to all disabled students, including blind and visually impaired students.
- There are benefits in using e-learning to complement traditional teaching approaches rather than on its own. This gives students the option of choosing the approach which is best suited to their particular learning style, as well as allowing them to obtain reinforcement by studying material using different approaches.

18.2.2 Educational Technologies, Devices and Tools

Some of the most important educational access technologies have already been discussed in Chapters 13 and 15 on computers and the Internet, and access to print media. Therefore, the content of this section largely comprises devices and tools that can be used for very specific applications, such as Braille and talking calculators and a talking dictionary.

18.2.2.1 Talking Dictionary

The Franklin 60000SE Language Master dictionary has more than 300,000 definitions and more than 500,000 synonyms and antonyms (words with similar and opposite meanings). The user can hear words and their spellings as well as phonetic pronunciations. The language master provides a dictionary, thesaurus and educational games. It has an adjustable voice speed and a large screen with adjustable type sizes. In the 'identify mode' keys speak their names when pressed without performing their functions. The device is hand-held and portable. It can be used with rechargeable batteries or a mains adaptor. Accessories include headphones, press-on locator dots, a carrying case, an instruction CD, a manual and a quick reference card. The Franklin Children's Talking Dictionary speaks words and definitions and shows how to write words. It includes a spelling list and interactive facilities for finding rhymes. It is hand-held and portable.

18.2.2.2 Moon Number Sheets

The Staffsmaths Moon number sheets from the RNIB use the Moon number system developed for visually impaired children with multiple impairments. The sheets contain several labels for each number and arithmetic sign and the labels can easily be removed from the backing sheet. They can be used in numerical work, to teach arithmetic and create worksheets, as well as to record phone numbers and label and adapt games.

18.2.2.3 Talking Calculators

Both basic and scientific talking calculators are available, with the scientific versions providing a greater range of different function keys. Cobolt Systems' talking calculator is a basic function calculator with percentage and memory keys. It has a large clear display, adjustable volume, large tactile keys and speech which can be turned on and off. When the earphone is plugged in the speaker is automatically disconnected. Their talking pocket calculator is a ten digit calculator with talking alarm clock. The date is displayed, but not spoken.

The Sci-Plus 300 talking scientific calculator form TechReady (UK) has a large, high contrast 8 digit LCD display with selectable speech output. The calculator has a large, high visibility keypad and colour differentiated operation keys. In addition to the four arithmetical operations, it has a range of scientific functions available. The calculator is supplied with universal recharger, headphones and large print user manual. The Institut de Réadaption de Quebec distributes audiocalc basic, scientific, and scientific and business. All three versions include stopwatch, alarm clock and calendar functions, a waterproof keyboard, earphone and case. The scientific version and, scientific and business versions are available in three languages, whereas the basic version is available only in French. In addition to scientific functions, the scientific and business version can calculate interest rate conversions, present and future values, correlation coefficients, linear regressions and cash flow computations. The talking scientific calculator MKII produced by the RNIB and the Open University has more than 50 functions in a sloping desktop case. The LCD (liquid crystal display) panel has 17 mm high characters and high contrast. The tactile keypad membrane has visual and tactile markings. The volume controls, headphone socket, charger socket and RS232 output ports are located at the back. It is supplied with stereo headphones and an RS232 serial cable for connection to a computer or printer, a charger and a carrying case.

The Accessible Graphing Calculator from ViewPlus Technologies was developed by the Science Access Project at Oregon State University, USA. This project is working on methods for making science, mathematics and engineering accessible to blind, visually impaired and dyslexic people. It is a software programme, which is compatible with most versions of Windows. It provides speech output and can be used by both blind and sighted people, as well as people with dyslexia. Its facilities include a scientific keypad calculator and an expression evaluator, which can import or define constants and expressions. The two data set pages can be used to compute expressions and standard statistical properties and import and edit data tables. A data set, its sum or difference and first derivative can be plotted, but more than one expression cannot be displayed on a graph and calculations with matrices cannot be performed. Graphs can be plotted on-screen or audibly using tones and the audio wave feature. The screen size can be magnified or reduced and the domain, range, use or lack of grid lines or tick marks and thickness of the graphlines can be adjusted. The user can adjust the speech rate, pitch and volume and many items can be selected with hot key shortcuts. Print copies can be made on a standard printer in a variety of fonts including Braille and print copies in Braille can then be copied onto swell paper and embossed. Alternatively, tactile graphics can be embossed directly onto a Tiger embosser.

18.2.2.4 Braille Calculator

The Leo Braille Display from Robotron Sensory Tools Division has a tactile 8-dot display and a Robotron mechatronic Braille cell. In addition to the four arithmetic functions, it has trigonometric functions, log, exp, ln, square root, powers, nested brackets and financial and unit conversion functions. It works on rechargeable nickel metal hydride batteries and a mains charger is supplied.

18.2.2.5 Rules, Metre Sticks, Protractors and Callipers

The RNIB distributes a plastic ruler with black markings and a measurement accuracy of 0.5 cm. The straight edge is for measuring and drawing and the notched edge to help placing pins when producing charts and graphs and using a pair of compasses. There are non-slip pads on the back.

The American Printing House for the Blind produces a number of measuring devices. The Braille/print protractor allows users to measure angles up to 180°. Large bold type numbers and Braille dots mark the degrees along the half circle of the protractor. The American Printing House 1 ft (30.5 cm) long brass ruler can be glued to sewing machines, drawing boards, band saws and other equipment. Raised lines indicate divisions down to 1/8 in (3 mm). Their metre stick is pre-drilled for hanging and has raised lines every centimetre and Braille every other centimetre. Their metric/imperial measurement ruler has both types of markings on different edges. One edge has raised lines at 0.5-cm intervals and Braille numbers at 2-cm intervals and the other raised lines at 0.25-in (0.6-cm) intervals and Braille numbers at 1-in (2.5-cm) intervals.

Spektra distribute a plastic calliper with tactile marks each 0.5 cm and folding 1- and 2-m rules with tactile marks every 1 cm. They also produce a set of plastic triangle, ruler and angle measures with tactile marks every 1 cm.

18.2.2.6 Braille Times Table

The Braille Times Table chart from the Braille Superstore is an 11 in (4.3 cm) square chart giving the multiplication tables up to 12×12 in Braille.

18.3 Employment

In considering the social integration of disabled people (and other minority or marginalized groups), access to employment is often considered to be of particular importance. There are a number of reasons for this, including the need for a good or at least an adequate income and the self-respect gained through being self-supporting. In addition, employment can provide a structure to the day and sense of purpose and often also a sense of identity, as well as providing the opportunity to make social contacts. For instance, a survey of young people with and without disabilities in the UK found that non-disabled young people had greater self-esteem and a slightly greater sense of control than young disabled people of the same age, but that the differences disappeared for young disabled people in employment (Hirst and Baldwin 1994).

However, employment rates for blind and visually impaired people are considerably lower than those for the sighted population, in most countries. Although employment and unemployment data is generally based on sample surveys of the labour force or households, it is frequently difficult to compare the data in different countries, due to different definitions of blindness, visual impairment, employment and unemployment. In addition, there may be different definitions of the working age and economic activity in different countries.

A survey for the European Blind Union performed in 2001 (EBU 2001), found that many of the 17 member organisations that responded were only able to provide estimates of the number of working age blind and visually impaired people, and that their unemployment rates were very high across Europe. Typical figures were 77% in Hungary, 87% in Poland, 72% in Germany, 55% in Finland and 68% in Norway. The two exceptions were Sweden and Spain with unemployment for blind and visually impaired people of only 5.5% and 4.2% respectively. However, in Spain 85% of the members of the national blind association ONCE who are in employment sell Lottery of the Blind tickets. Clearly, this type of employment does not have a career structure with promotion prospects and it is unlikely that it is matched to the skills and interests of 85% of blind people. Unfortunately in the employment area, governments and other organisations have tended to focus on getting disabled people into employment of any type and give little attention to issues of job satisfaction, career structure and promotion. Barriers to the employment of blind and visually impaired people include (EBU 2001) a lack of qualifications and experience, high general unemployment, employers' prejudices and poor legislation. Barriers to employment found in an Australian survey include discouragement, discrimination and lack of information about suitable career choices (Smith 2002).

Comparative statistics for visually impaired and non-disabled people or the total population are available in a number of countries, though not all the data is recent. For instance, in the UK in 2001, some 44.3% of blind and visually impaired people were in employment compared to 78.9% of the total population or 85.3% of the non-disabled population of working age (16–64 for men and 16–59 for women) (Smith A and Twomey 2002). In the USA, some 30% of legally blind people and 40–45% of blind and visually impaired people compared to 82% of the non-disabled population of working age (18–69 years) were in employment in 1994/1995 (Kirchner *et al.* 1999).

An Australian survey of blind and visually impaired people identified the critical role played by assistive technology, with 21 of 25 respondents using some form of assistive technology, particularly to access computers and the Internet. Respondents who were required to travel as part of the job were confident about travelling independently (Smith 2002), indicating that improved orientation and mobility training and improved confidence in this area may overcome this barrier to employment for some blind and visually impaired people. Although just over two-thirds of the respondents believed that they had been treated fairly with regards to career development, the remaining just under one-third considered they had experienced discrimination and that there was a 'glass ceiling' or barrier to further advancement for visually impaired people (Smith 2002).

18.3.1 Professional and Person-centred

This category covers activities in a wide range of different employment fields, including the health care professions, therapy and counselling, the legal profession, management and personnel, accountancy, banking, religious organisations, translation and interpretation, journalism and creative writing, acting, dancing, music and social and community work. The main requirements for modified or assistive technology in the workplace are related to information and communications technologies, particularly access to print media and computers and the Internet. There are a number of professional associations to support blind and visually impaired people working in these areas, particularly in the USA and some of them are listed in the Resources section at the end of this chapter. For instance, the National Association of Blind Lawyers provides support and information on employment, techniques used by blind people and laws affecting them, as well as advocacy.

Owing to the increasing availability of assistive technology for accessing print and computers in many countries, one of the main remaining barriers to blind and visually impaired people obtaining professional of person-centred employment is likely to be attitudinal, including expectations about suitable employment for them, rather than practical. For historical reasons, physiotherapy is one of the professions in which blind and visually impaired people are best represented and a brief case study is given below. Music has been another popular profession amongst blind people and there is some research evidence that musicians who were born blind or become blind early in life are better at distinguishing the pitch of notes, particularly those that are close in pitch or of very short duration, than sighted musicians or those who became blind later in life (Gougoux *et al.* 2004).

18.3.1.1 Physiotherapy

Physiotherapy has been a significant profession for blind and visually impaired people for a number of years, at least in part due to its historical relationship with massage (French 1995). In Japan, massage was practiced solely by visually impaired people for several hundred years (Yoshimoto 1901). In the UK, the Association of Blind Chartered Masseurs was set up and became the Association of Blind Chartered Physiotherapists (ABCP) in 1953. There were about 50 certified visually impaired masseurs in the early 1900s and more than 900 visually impaired people, including 200 ex-servicemen blinded in action, had trained as physiotherapists through the RNIB by 1987. However, there was some opposition to visually impaired physiotherapists from the profession, partly on gender grounds, as many of the early visually impaired physiotherapists were male (ex-servicemen), whereas the rest of the profession was female. Visually impaired physiotherapists also threatened professional status and distance from the people being treated by being disabled and often being of working class rather than middle class backgrounds. The NIB (which became the RNIB in 1953) included medical electricity in its syllabus in 1919, but the professional body only agreed to examine electrical treatments other than ultra violet light in 1934.

With the development of the profession, technology became more important and RNIB scientists developed a number of Braille and audible devices for use in the profession. The Braille galvanometer (the historical name for a moving coil electric current detector) was a device that allowed blind and visually impaired people to carry out treatment with continuous direct and interrupted alternating current (then called galvanic and faradic current respectively). An audible device was developed to tune the short-wave diathermy machines used to induce heat electrically and give muscle relaxation. An erythemameter, which used audible signals to indicate the redness of the skin, was developed in the early 1950s and this allowed blind physiotherapists to use ultra violet light in treatment. However, in practice, few physiotherapists used the erythemameter and some considered it was developed for political reasons and would only be required if both the physiotherapist and client were blind and totally isolated from sighted colleagues. The statement that visually impaired physiotherapists had met the requirements 'in accordance with the syllabus for blind candidates' was dropped in 1975 and the identification of visual impairment on the Chartered Society of Physiotherapy (CSP) register ceased in 1988. Blind physiotherapists now face an increasing number of barriers, due to the inflexible environments and structures in which they work rather than changes in the profession, as a result of which a high proportion of visually impaired physiotherapists practice privately from home (French 1995).

18.3.1.2 Theatre

There are theatre companies of blind people in Croatia, the USA, the UK, Slovenia and Spain. The oldest is New Life in Zagreb which was set up in 1946 and had its first performance in 1948. It has toured all over Europe and organised the First International Blind and Visually Impaired Theatre Festival in 1999. Theatre for the Blind was set up in New York in 1979, originally as a sighted company. Each actor receives a version of the script in an appropriate format on the first day of rehearsals. This includes scripts in Braille and on tape, as well as very large print (40 pt). Rehearsals last longer than those of sighted companies to give performers additional time to familiarise themselves with the script and stage set. The designers create stage environments which are helpful to blind actors. For instance, one production had a tiny almost invisible rim at the front edge of the stage, to inform actors where the playing area ended. High contrast colour schemes are also used to aid visually impaired actors.

18.3.2 Scientific and Technical

This category includes employment activities in engineering, technology, computing science, programming and information technology, mathematics (other than accountancy and banking) and the natural, physical, social and human sciences. In addition to access to information technology and print, activities in this area often involve laboratory work and accessible laboratories are the focus of this section.

18.3.2.1 Laboratory Work

Both education and employment in engineering and the physical sciences, amongst other areas, require practical or laboratory work. In the educational context it is important that laboratories are fully accessible to both disabled students and staff, including blind and visually impaired students and staff, so that students can participate in the same educational experiences as other students and staff are able to carry out teaching and research effectively. In the employment context laboratory accessibility widens the range of career opportunities open to blind and visually impaired, as well as other disabled, people.

Although different types of laboratory involve very different types of equipment, the following principles of accessibility, most of which have been derived from a mini-laboratory accessibility project (Hersh *et al.* 2004) are relevant in all cases:

- Adequate and appropriate lighting, an uncluttered layout of the laboratory and designated places for all equipment, to make it easier for visually impaired and blind students to find the equipment they want.
- As far as possible, using only one model of each type of equipment, for example, only one type of oscilloscope from the many types that are available. This means that for each type of equipment blind people only have to familiarise themselves with the layout of one rather than multiple devices.

- Use of dymotape Braille and/or dymotape with embossed large print letters to mark equipment.
- Use of equipment with audio output, RS232, USB (universal serial bus) and/or other computer connections to enable the transmission of data to a computer for use with screen readers.
- Ensuring that all specialised software that might be used in laboratories, such as LabView™ and MATLAB®, is fully compatible with screenreaders.
- The provision of supports, particularly for fragile equipment such as glass vessels, to reduce the likelihood of breakage and injury to blind or visually impaired laboratory users.
- The provision of all equipment documentation and laboratory sheets in electronic format (with appropriate descriptions and explanations of any graphics) so they can be accessed *via* screenreaders or a Braille display.
- Avoiding curriculum aims which disadvantage blind and visually impaired students, such as the requirement to be able to draw graphs manually rather than the requirement to be able to input data to appropriate software, use the software to produce graphs in appropriate formats and analyse the resulting graphs.

Some talking equipment is available, but further developments in this area are required. It should be noted that talking laboratory equipment is also useful to people with dyslexia and may also have some benefits for sighted people, particularly if there is the option of customisation to choose the output modalities.

18.3.3 Administrative and Secretarial

The coverage of this activity category is approximately equivalent to the membership categories of the US National Association of Blind Office Professionals, which comprises blind secretaries and transcribers, including medical and paralegal transcribers, office workers and customer service personnel. The Association addresses technology, accommodation, career planning and job training. Access to information technology, print media and telecommunications is very important for activities in this category and the associated assistive technologies were discussed in Chapters 12, 13 and 15. Some technology to access liquid crystal displays (LCDs) is considered next in this section.

18.3.3.1 Accessing Liquid Crystal Displays on Equipment

An ever-increasing number of products, including telephones, fax machines, photocopiers, modems, test instruments and bar code readers, have text liquid crystal displays (LCD). It is therefore important that blind and visually impaired people can access these displays.

The LCD interface developed by the Kentucky Department for the Blind in 1993, whose schematic is shown in Figure 18.3, gives blind and visually impaired people access to an LCD display. It can be used with all products that have an LCD that

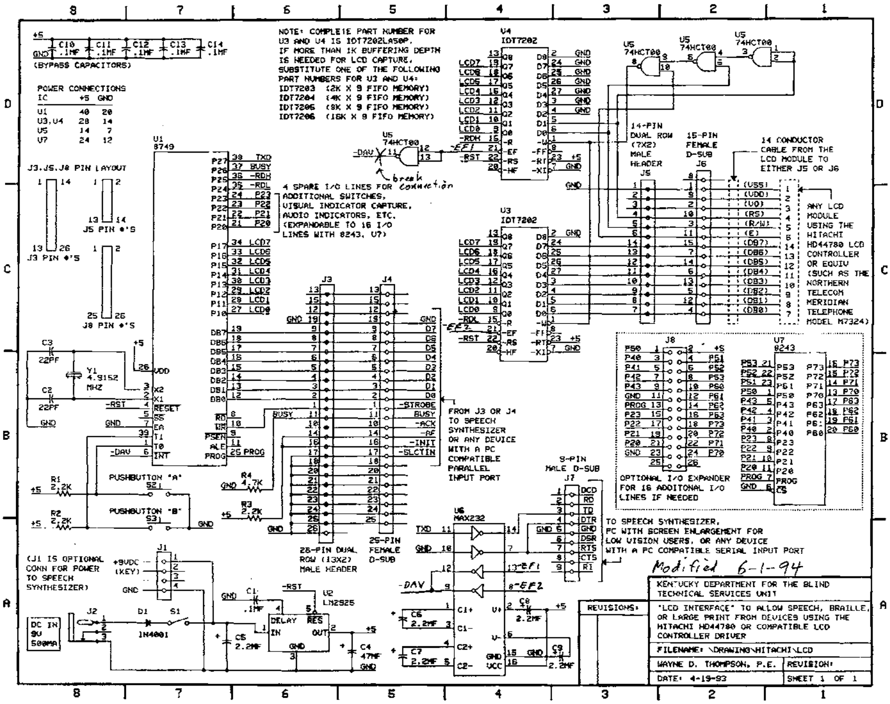


Figure 18.3. Schematic diagram of LCD interface (photograph reproduced by kind permission of Mr. Wayne Thompson (designer) and the Kentucky Office for the Blind, USA)

is compatible with the Hitachi DH44780 LCD controller. Since this controller has become almost a *de facto* standard, this is not a significant restriction.

The LCD interface circuit is connected to the device's 14-pin bus, which is used by the HD44780 compatible LCD controller. This allows the LCD interface to capture copies of all the information written to the LCD. The capture process takes place automatically in real-time and does not affect normal LCD operations. The process is transparent to the product's normal operations.

The connection process can require a soldered connection to be made inside the product or simply the user to plug a cable into a built-in externally accessible connector, depending on the extent to which the manufacturer has complied with the proposed LCD interface being considered by the National Federation of the Blind. After capture the LCD data is processed by the interface and sent to the on-board parallel or serial output ports. The data can then be presented to the user using a variety of existing speech, Braille and character enlargement devices. Customisation of the LCD interface can be used to present the output in the desired format or combination of formats for use with built-in hardware or an appropriate assistive device.

The interface adaptation involves the following three main steps:

1. Connecting a cable to the product's 14-pin LCD module using a J5 or J6 connector, according to whether the interface is embedded in the product or housed externally.
2. Constructing the LCD Interface.
3. If necessary, modifying the LCD Interface firmware to best fit the product's operational characteristics.

18.3.4 Skilled and Non-skilled (Manual) Trades

Employment categories covered by this activity category include the building and decorating trades, shop work, industrial and factory work, janitorial work, the catering and hotel industry, security and monitoring and the police. Many of the activities in this area involve access to machinery and equipment. Some of the devices and tools with tactile or audio output and the assistive technology for providing readouts from inaccessible equipment and machinery are discussed in this section. However, it should be noted that the range of tools, equipment and machinery available to blind people is still very limited and most blind and visually impaired people may be unaware of what is available.

18.3.4.1 Using Machinery and Equipment

Skilled and unskilled (manual) trades involve a wide range of machinery and equipment. It is most likely that existing workplace health and safety legislation and regulations were derived on the assumption that all employees will be non-disabled. Consequently, there will be a need to extend this legislation and the concomitant regulations and guidelines on a *design for all* basis to take account of the presence in the workplace of visually impaired, blind and other disabled people. It is also essential that all employers have adequate procedures in place for implementing health and safety legislation and regulations, and that compliance is strictly enforced. However, it is recognised that there may be hazardous industrial or other workplace situations where appropriate safety procedures cannot be derived for either blind, or visually impaired or some other groups of disabled people.

DigiCite from Compusult is a voice output digital display reader that constantly monitors and captures images of digital displays. Whenever the image changes it is converted to text and notification given using recorded digitised speech or other optional outputs. It can be used with telephone consoles, alarm displays, measuring devices, recording and testing equipment and electronic devices. Since it uses digitised speech rather than voice synthesis, it can accommodate any language and even several different languages in one system. It consists of a small computer, an image capture device, mounting hardware and a power supply.

Responder from Compusult is a voice output monitoring system, which uses machine vision and speech technologies to provide an automated monitoring, and notification system for workplace and industrial installations (see Figure 18.4). This

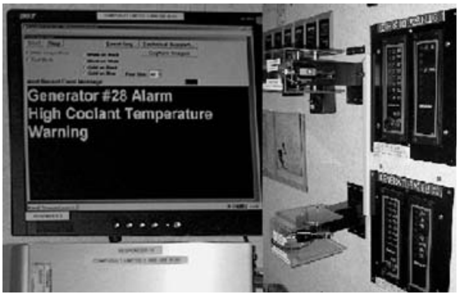


Figure 18.4. Responder monitoring system (photograph reproduced by kind permission of Compusult Ltd., Canada)

enables blind and visually impaired people to work in security and monitoring environments. The system continuously watches electronic equipment, such as alarm panels, and provides notification *via* recorded and/or synthetic speech over loudspeakers, public address systems, telephone systems and/or modems. It also keeps a complete log of all events to help in verifying the operational status of the equipment. The system is unobtrusive and does not require modification of existing alarm panels. It is able to respond to light and sound indicators, alphanumeric digital displays and analog and digital output from other systems.

It can be configured to meet particular requirements. There are two main configurations. Responder M1 comprises a single computer and one to three black and white digital cameras. It is particularly suited to environments where notification is only required for on/off type alarms and events. Responder C1 is able to distinguish between different colours. It can therefore provide notification of alarms or events indicated by colours, for instance, green, yellow and red situations. The two Responders M1 and C1 can be networked together or deployed separately. They can be used on IBM and compatible computers and require Windows NT 4.0 or above.

Vindicator from Compusult is a speech enabled LED monitor that uses electrical and optical sensors to detect on/off events and to provide notification to blind, visually impaired, hearing impaired and deaf people. The standard output is digitised speech in any language. Optional outputs include large format alphanumeric displays, LED indicators, audible tones, serial interfaces to computer systems and electrical signals. Vindicators are transportable and do not require internal connections to existing electronic equipment.

The Vindicator Standard detects when LED, VCF or LCD display indicators turn on and off and provides notification using digitised speech or other optional outputs. Different words or phrases can be used to indicate the type of indicator and whether it is switching on or off. Vindicator Standard is available in 8 and 16 line versions, with one line for each electronic display indicator. There are both speech enabled stand-alone units and units with a serial interface to computer systems and other electronic equipment. The serial interface is compatible with Compusult's DigitCite and Responder systems. The Vindicator Wired detects voltage on/off conditions on electrical panels and other devices and provides notification in digitised speech, large format LED displays or an encoded output *via* a serial interface to a computer system. It is available in 8, 16 and 40 line versions, as well as stand-alone and serial interface versions. The Vindicator RF (radio frequency) Switch detects when equipment, switches or indicators are on or off and transmits notification *via* radio frequency signals to a receiving unit for output using audible tones and large LEDs or digitised speech. It is used to identify events away from the user's general work area and does not require external wiring between locations. It is available in single or multiple transmitter/receiver versions. Vindicator Buttons detect when an LED, VCF or LCD display indicator switches on and off and provides notification using audible tones. It is a small stand-alone device designed to fit over a single electronic indicator to provide audible output. Vindicator Buttons have various audible tones to help identify different indicators and events.

WorkCam from Compusult provides optical and digital magnification of equipment and materials to enable visually impaired employees to carry out hands-on work with small-scale components. It uses a digital camera mounted on a flexible or articulating arm to interface with a video monitor or computer system to provide a live high resolution high magnification display. WorkCam is available with colour or monochrome video output and a number of mounting, magnification and other options. Automatic digital camera zoom, focus and iris controls can also be provided.

18.3.4.2 Battery Tester

CareTec distribute a battery tester that can be used for 1.5- and 1.9-V batteries and rechargeable accumulators. The states of completely charged, half power and running down are indicated by three, two and one audio signals respectively. The US National Federation for the Blind distributes an audible battery tester that is able to test 9-V, AA, AAA, C, D and N size batteries. The battery is placed in the tester and a button pressed. A continuous beep is heard if the battery has power and its loudness increases with the amount of power remaining. There is also a light indicator.

Ann Morris Enterprises' audible batter tester (model ABT2) is a hand-held battery tester that can test AAA, AA, C, D, N, 9-V and button cell batteries. It emits a strong buzz, a weaker buzz and no sound respectively when the battery is charged, marginal and discharged. There is also a visual display with green, red and yellow zones.

18.3.4.3 Talking Tyre Gauge

The *accutire* talking tyre gauge displays tyre pressures on an LCD screen and reads them aloud. The recall button allows the reading to be repeated. The gauge reads from 5 to 99 pounds per square inch (PSI) (34.5 to 682.6 kPa) with an accuracy of $\pm 1\%$ at 0.5 PSI (3.5 kPa). It should be used with cool tyres for precise measurements. The Bilingual talking tyre gauge from Radio Shack reads tyre pressures in either English or Spanish and also has a large LCD display. The recall feature repeats the last reading. It can measure pressures from 5 to 99 PSI (34.5–682.6 kPa). The talking tyre gauge from Brookstone displays and speaks tyre pressures up to 150 PSI (1034 kPa).

18.3.4.4 Spirit Level

Seiko's spirit level can be used to check that surfaces are either horizontal or vertical. It can be used with green and red LED lamps or with lamps and audible tones to indicate when the device is level. High and low tones indicate that the left and right side respectively are high and no tone that the unit is level. A ruler can be pivoted to any angle and locked in place and can check vertical surfaces when it is set to 90°. A V-shaped channel in the bottom of the device can be used to locate it on curved edges such as pipes.

18.3.4.5 Adapted Tools

SpeakFAST from Compusult Ltd (Canada) provides access to the FAST Technologies Sensor System 1000 torque wrench for blind and visually impaired people (see Figure 18.5). It captures data from the torque wrench and converts it to speech output or torque readings. The two units can be connected together into one unit. SpeakFAST can be used with headphones or amplified speakers. The FAST Technology torque wrench converts a wrench into a precision torque setting system and provides instantaneous measurements by a digital meter with an LCD display. It has an acoustic alarm. The measurement range is ±150 Nm with a precision of <3% and there is a peak value hold.

The measuring tool with speech output from Brailletec was designed for connection to the Voltcraft multimeter series, as well as the Mitutoyo slide calliper. The multimeter can be used for measuring voltage, current, frequency and other electrical quantities. The unit has a serial data port for processing data on a PC. Speech output is activated by a push button and is deactivated after the announcement to reduce the likelihood of accidental discharging of the battery. The left DIN

Figure 18.5. SpeakFAST (photograph reproduced by kind permission of Compusult Ltd., Canada)

socket should be connected to the PC and the charger plug for the internal battery. The right socket should be connected to the slide calliper and the charger plug for the internal battery. The slide calliper and multimeter should not be operated simultaneously.

18.3.5 Working Outside

This employment activity covers agriculture in all its diversity, gardening, estate management, landscape design, working in zoos, horse riding and horse racing. In the 'developing' countries, blind and visually impaired people may have little choice about engaging in (subsistence) farming in order to survive. In the industrialised countries, these areas may not be the first employment choices for blind and visually impaired people and there is little information about either assistive technology or human support. However, the US National Federation of the Blind does have an Agriculture and Equestrian Division.

18.4 Recreational Activities

Recreational activities cover a wide range of leisure, sport, entertainment and social activities that take place in a range of different locations. A full categorisation of recreational activities is obtained from the third level of the CAT model, as is shown in Figure 18.6. This is used in this section to systematically detail the assistive technology solutions available to overcome barriers to independent participation in recreational activities. It should be noted that blind and visually impaired people do not generally require additional assistive technology or technological adaptations for activities in the friendship and relationship category to those required for the fundamental activities of mobility and information and communications. However, some blind children may require additional help in developing social skills, which sighted children learn at least in part by visual observation.

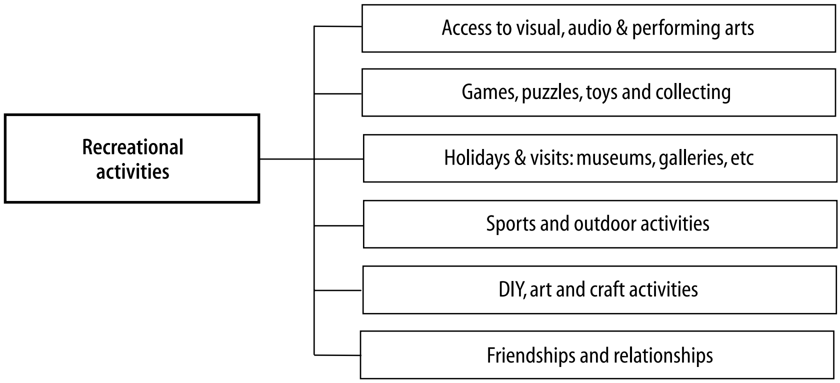


Figure 18.6. Recreational activities: the third level areas

18.4.1 Accessing the Visual, Audio and Performing Arts

This section considers access to television, video, theatre, cinema, radio, concerts, opera and dance performances. One of the main solutions to making theatre, television and cinema accessible to blind and visually impaired people is audio description, which is discussed in Section 18.4.1.3. The principle of a combination of an audio description provided by a human expert with transmission of the description by technology to the viewer is the same in all cases, though the details of the technological part of the solution vary with the application. In the case of television, another solution is offered by screen magnification. Unlike cinema and theatre, television, video recorders and related appliances need to be controlled by the viewer. Therefore, this section also includes discussion of accessible remote control devices for television and related appliances.

Listening to music does not in itself pose problems to blind and visually impaired people without hearing impairments. However, they may encounter barriers in operating the controls of a radio or accessing programme information provided at a concert or opera. The solutions presented in this section are talking radios and audio introduction.

18.4.1.1 Accessible Remote Control Devices for Television and Related Appliances

The talking video remote control from Cobolt Systems is compatible with a range of standard video recorders. It can be used as a remote control or programmed to record programmes at set times. It is then left in the same room as the recorder and will send appropriate signals to turn the recorder on and off at the desired times.

The Oversized Universal Remote can control up to four devices, including a television, video recorder, DVD (digital versatile or videodisc) player and cable box or satellite unit. The buttons are very large and the characters are bold black on white buttons. The keys light up for easier recognition in a dark room. The remote is also large to ensure it will not be lost.

Guidelines for accessible remote control devices include the following (Petré *et al.* 2006):

- Large, well separated buttons with the spacing between buttons at least half the width of the button.
- Keys should be raised or have raised edges to make them tactually discernable and should have distinct shapes or textures to help distinguish between functions. Tactile indications should use international standards.
- The buttons should be logically grouped and spaced for easy manipulation. Smaller spaces should be used between groups of buttons relating to the same function than between those for different functions.
- There should be a raised marking on the number 5 key on the numeric pad.
- Clear visual markings with any text or symbols on the remote buttons large enough to be clear, legible and of a contrasting colour to the keys or background, as well as durable.

- Labels should be clear, intuitive and standardised, with clarity about the meaning of options such as ‘exit’ and ‘save’.
- Audible and visual feedback should be provided when a remote control key is pressed with the option to turn off this feedback.

18.4.1.2 TV Screen Magnification

Fresnel lenses, such as the MagnaScreen, can be put in front of a television screen to magnify the picture. The magnification that can be achieved without unacceptable distortion and loss of contrast is relatively low at about 3×. Fresnel lenses are flat on one side and ridged on the other. Plastic Fresnel lenses give a poorer quality image when compared to that of a continuous glass lens. Consequently, plastic Fresnel lenses are used as magnifiers in applications where a thin light lens, but not perfect image quality, is required. They consist of a large number of concentric rings stacked together. Each ring is slighter thinner than the next one and focuses the light towards the centre. The angle of each ring’s angled face is different to give the whole lens the ability to focus light toward the centre. TV Zoom zooms in on part of the TV picture, giving up to 4× magnification. It takes RF and SCART (Syndicat des Constructeurs d’Appareils Radiorécepteurs et Téléviseurs) input and is connected to a TV *via* SCART. It is compatible with analogue and digital television, but not analogue teletext. The Topolino television combines TV zoom and a video magnifier that can magnify print up to 40×.

18.4.1.3 Audio Description of Programmes and Performances

Audio description is the descriptive narration of important visual elements of theatre, television, film and other media. It provides audio explanations and descriptions of settings, characters and action that are not given in the regular audio presentation. In theatres and similar venues, the user receives the audio description *via* a small earpiece or headphone connected to a tiny receiver. The aim of audio description is to give blind and visually impaired people a fuller understanding of what is being shown and the same sort of appreciation of the presentation as a sighted person.

Good description should be clear and not confuse, mislead or distract from the soundtrack. The language used should be simple, clear, descriptive, precise and accurate. Offensive terms, terms indicative of prejudice and technical terms should be avoided as far as possible. The description should not override the dialogue, so it does not interfere with the performance, unless the information provided is essential, for instance if someone draws a gun. The description should totally avoid interpretation and personal comment and the style, tone and pace should be harmonised to the content and mood. Different styles of description will be appropriate to different types of presentations. Audio description includes descriptions of action sequences, facial expressions, costume, and the scenery (ITC 2000).

In live performances, a trained audio describer transmits the descriptions from a soundproof booth in the theatre, generally using a court stenographer’s mask microphone. The description is relayed to a headset or earpiece worn by the theatregoer from the soundproof booth *via* an infrared or frequency modulated (FM) transmission system. Headsets or earpieces are generally available from the theatre box office. Audio description provides a verbal commentary of what is happening on the stage for blind and visually impaired theatregoers.

In films, the description is generally pre-recorded and synchronised to the sound track. In television, the audio description is transmitted to the user *via* the second audio program or SAP channel on their TV or videocassette recorder (VCR). TVs and VCRs manufactured from the early 1990s on are generally able to receive SAP signals. Audio description is available on digital but not analogue television, which means it is not available to everyone. For instance, in the UK currently about one-fifth of the country cannot receive digital television, though this will change as analogue television is phased out. In DVD videos, the audio description is accessible *via* the DVD menu as a special feature of the languages menu.

Commercially available computer-based workstations with software for work processing, videotape control and digital audio editing are now used to produce audio descriptions for television. The work station generally includes a personal computer, a time code index, a video edit controller, a prompting device for recording the description in the gaps in programme dialogue, a time-coded VHS (video home system) or DVD player, a small monitor with associated loud speakers and a device to store the descriptive audio. The PC should be able to associate the elements of the written script with the programme time-code. When the audio description is recorded, the loudness of the narrative voice is fixed at the start of the recording, but the background level can be adjusted, for instance to reduce loud programme sounds or audience laughter. Continuous traffic noise or laughter throughout several consecutive descriptions should be kept faded to a lower level so that the full background volume does not come through between descriptions (ITC 2000).

To receive digital television, viewers require a television adapter, an integrated digital television or a computer adapter to allow them to view television on the computer. However, a specific set-top box and computer adapter are required to obtain audio description. i-Player Plus from Netgem has a different architecture from other set-top boxes that enables it to receive audio description. It can be connected with a SCART cable to a TV or video recorder. Connection to the telephone line with the 8-m cable provided will allow the software to be updated and additional services provided. The system includes the set-top box, a remote control, two SCART sockets for connection to the TV and video recorder, S/PDIF (Sony/Philips Digital Interface Format) optical audio output, a 3.5-mm stereo jack socket audio output, a USB1.1 socket and a telephone socket, as well as the associated leads and cables. Instructions are available in large print and on an audio CD. A built-in voice guides the user through the installation and informs them of the channel selected. The RF (radio frequency) loop-through does not include an output from the digital receiver and therefore the SCART socket is required and the TV and video recorder cannot be tuned to the set-top box. An RF modulator can be inserted between the i-Player and TV aerial if the TV and video recorder do not have a SCART socket.

The Portset Digital Media Centre is a stand-alone TV access device without a screen, which provides sound-only digital terrestrial television with audio description, access to electronic programme guides, a Daisy talking book player and a recording function. It incorporates the Nebula Electronics DigiTV digital TV receiver, and is designed to make television, radio and teletext accessible to blind and visually impaired people. Programme information can be accessed through the integral keyboard or remote control and is provided through speech messages to the user. The system uses the Nebula Electronics audio description module to describe the scene shown on the TV. The system is also able to digitally record television and radio programmes and includes an integrated music CD, MP3 (Moving Picture Experts Group Audio Layer 3) and Daisy talking book player, as well as a fully accessible DVD player that plays audio description content.

The major and many of the smaller, distributors are making the majority of their films accessible. About a quarter of UK cinemas have installed the equipment necessary to screen these films. Yourlocalcinema.com is a listings service of accessible (to visually or hearing impaired people) film showings in the UK and Ireland. It provides a staffed UK call centre, a web over phone service and a talking website for visually impaired people, which talks without requiring screen reader software.

Cinemas broadcast audio description through infrared headphones that can be collected at the box office on arrival. The film soundtrack is broadcast as normal through the cinema's surround sound speakers and the audio description soundtrack describes the on-screen action at gaps in the dialogue. Special showings of films with audio descriptions are not required and only cinemagoers with infrared headphones are aware of the description.

The UK Independent Television Commission has produced guidance for audio description based on studies carried out by the European Audetel (Audio Described Television) consortium, between 1992 and 1995 (ITC 2000).

Audio description can also be used to describe paintings and photographs and make them accessible to blind people. Audio description is normally intended to be as neutral as possible in order to provide blind and visually impaired people the same type of information as is available to sighted people. However, it may not be possible to describe adequately a painting or other work of art while remaining totally neutral and objective.

Thus, radio, cinema and TV all use the same approach of audio description, with slight differences in when the description is produced and how it is transmitted to the viewer or audience member. The use of audio description also illustrates a good balance between the use of technology and the use of human assistance, in this case through a qualified audio describer. Technology is used to transmit the audio description to the viewer or audience member, but the actual description is produced by a person, as human input is required to determine what details are relevant and interesting to viewers and present them in a way that enhances rather than detracts from the whole performance.

18.4.1.4 Radio Audio Description System

The radio audio description (RAD) is a system used for audio description in large venues, which transmits on newly freed up sections of the broadcast band radio spectrum. An RAD transmitter is professionally installed for permanent location in each venue and staff are trained in its use. Each system can be used by two commentators, each with their own microphone headsets, at any one time. A line level input can be used to send pre-recorded information before the start of the performance to enable users to check receivers and adjust reception levels for comfortable listening. RAD receivers are small lightweight radios with earpieces and disposable batteries. They have only two controls: a combined on/off switch and volume control, and a fine-tuning control, making them simple to operate. Unlike infrared or low power radio receivers they do not need to be in line of sight or close to the transmitter. Each transmitter can send to an unlimited number of receivers. Receivers and earpieces can be stored in flight cases for 10, 20 or 50 units. Since RAD system operate in broadcast band frequencies, they need to be licensed by the relevant national regulatory body.

18.4.1.5 Talking Digital Audio Broadcasting (DAB) Radio

PURE Digital, a division of Imagination Technologies, a UK company has produced the SONUS-1^{XT} DAB digital radio with iVOX voice-feedback technology and voice prompts, as well as a clear graphical display and alarm clock functions with a voice output (see Figure 18.7). Tapping the handle on the radio twice confirms the alarm settings. The digital radio talks the user through all stages of setting up and tuning and speaks the station names as well as displaying scrolling text. There are ten presets and automatic volume equalisation means that the user may not need to adjust the volume when changing stations. The time and alarm settings are announced when the 'snooze handle' is tapped once or twice respectively. The clock is fully automatic and updates itself from the radio signal, including changing between summer and winter time settings. There are nine record timers connected to an optical output to allow the user to record radio programmes to a separate recording device. This requires the use of a recorder with an optical input that supports sync-record, such as MiniDisc or CD-R recorder. The radio has a USB connector for updating the device over the Internet, optical digital out, stereo line out, headphone output, XT-1 speaker output and power input. It is currently only available with English speech and pre-programmed with the names of UK radio stations, but other options are being considered. The price is reasonable for a DAB digital radio. Stereo output can be obtained by connection to the optional XT-1 speaker.

18.4.1.6 Audio Introduction

Some opera houses and concert venues have introduced audio introduction, which is similar to audio description. It is used in opera houses before the start of the performance to provide a resume of the story, a description of the set and costumes

Figure 18.7. SONUS-1^{XT} DAB digital radio (photograph reproduced by kind permission of PURE Digital, Imagination Technologies, UK)

and a clear indication of the order in which the singers appear. In concert halls audio introduction is used to present the print programme together with information about forthcoming events. Music scores for blind and visually impaired people, including the use of Braille and spoken music, are discussed in Chapter 16.

18.4.2 Games, Puzzles, Toys and Collecting

A range of games, puzzles, toys and collecting activities are popular with both children and adults. Most of the popular board and card games are available in fully tactile or large print versions, as discussed in this section. Talking computer games have also been developed. There are also issues of whether or not other computer games are compatible with screenreaders, but the authors are not aware of any surveys.

18.4.2.1 Games

Most games rely on visual feedback of the status of the game and in their original format, are consequently inaccessible to blind and visually impaired people. Therefore, making games accessible requires feedback on the status of the game to be available in tactile or audio format for blind people and large print high contrast format for visually impaired people. Due to the relative costs of providing speech or tactile output (Braille, raised or engraved lettering, game pieces of different sizes and shapes and/or holes in the board), most games designed for blind and visually impaired people provide tactile rather than audio feedback. This has the advantage of making them accessible to deafblind people.

Examples include the following from the RNIB online shop:

- Draughts, with the lighter coloured squares on the board recessed to distinguish them by touch from the darker coloured squares, large light and small dark coloured wooden pieces and two easy to see tactile dice.

- Snakes and ladders, with playing pegs of different shapes and colours, different colours for squares, snakes and ladders and a shape/colour system which allows the game to be played by touch and/or sight.
- Chess, with a wooden board and wooden pieces distinguishable by touch, as well as a mechanical chess clock with tactile markings and an audible signal when time is up.
- Monopoly with a tactile board, fully labelled in Braille, tactile dice and all title deeds and cards labelled in large print (16 pt) and Braille).
- Wider than normal playing cards to make them easier to see and use, with large print and Braille markings.

There are also tactile books, blocks and puzzles, as well as other toys designed for blind babies and very young blind children. The RNIB also produces a set of 12 scented water colour markers, with the colour identifiable by the scent.

18.4.2.2 Computer Games

Computer games are becoming increasingly popular. GamesForTheBlind.com produces a number of talking games for blind and visually impaired people. The games include a copy of the Eloquence speech synthesiser, which has eight different voices, covering both sexes and different ages. The pitch, speech rate and volume can also be modified to create a voice that is easy to understand and that speaks at an appropriate rate. The games can be played on a computer using versions of Windows from 95 upwards and does not require any additional hardware and software (other than computer speakers). The initials SV after the name of the games indicate that they are talking or self-voicing.

18.4.3 Holidays and Visits: Museums, Galleries and Heritage Sites

Going on holiday generally involves accessing information and mobility activities, the associated technology for which is discussed in Chapters 5–16. Another important aspect of holidays is sightseeing and visiting museums, galleries and heritage sites. Although one of the earliest touch tours was in 1976, it is only relatively recently that museums, galleries and heritage sites have started to think seriously about blind and visually impaired visitors.

18.4.3.1 Accessible Museum Visits

Tactile exhibitions, often referred to as the ‘touch scene’, in museums are aimed at three different audiences: children on school visits, outreach work with elderly people and visually impaired visitors. One of the earliest touch tours, in which pre-selected objects could be handled, was produced by the Tate Gallery in London in 1976. Unfortunately, visually impaired visitors to this exhibition were not allowed to take sighted friends along as guides and consequently had problems finding their way around and finding the objects to be handled. The first permanent touch tours in a UK museum were introduced into the British Museum in 1990. Braille and large print signs indicated that certain sculptures in the Egyptian and Graeco-Roman galleries could be touched. Sighted people, including accompanying people, were not allowed to touch the designated sculptures. Museum guides could be provided for unaccompanied visually impaired visitors if sufficient notice was given. Somewhat ironically in view of the claims that it should be returned to Athens, it is the Parthenon Frieze that has been the recent focus of the British Museum's accessibility work. The new accessible gallery includes a plastic moulded section of the frieze, an audio guide and a book with a series of simplified raised diagrams. The diagrams show both tactilely and visually what is depicted in the frieze. There is a Braille key and a series of tactile symbols to help with identification, as well as a 5-h taped commentary to guide in using the book (Hetherington 2003). However, the book is a simplified visual depiction that can be touched rather than a source of haptic access.

Tactile versions of a selection of photographs at the Natural History Museum have been made from cellulose acetate, with layered black and white scans of the original photos, interpreted in relief using precision etching techniques. The different textures are used to suggest buildings, animals and landscapes. Some of the tactile pictures require the content to be supplemented by text description, whereas others are immediately comprehensible. Unfortunately, tactile versions are not available for all exhibitions.

Many museums and galleries have audio guides, some of which have been designed specially for blind and visually impaired people. Some museums have audio or scents as part of their display. Some venues produce an access guide with information on services and facilities for disabled visitors. An increasing number of museums, galleries and heritage sites organise a programme of events and workshops, which may include touch tours and events at which the contents of the venue are described. English Heritage produces a guide that indicates which venues have audio commentaries and allow sculptures and stonework to be touched. Its venues provide water bowls for guide dogs on request.

Additional information on haptic access to museums and art objects can be found in Chapter 4. For instance, The Museum of Pure-Form is a European Union Project to make virtual copies of museum exhibits available for haptic exploration (Jansson *et al.* 2003).

18.4.4 Sports and Outdoor Activities

The ability to participate in sport is important both for social networking and as a means of exercise and maintaining health. Many sports, such as football, tennis and archery, use visual information and therefore visually impaired and blind people require support from assistive technology or a sighted assistant to play them. Despite the potential for the development of assistive devices in this area, there are few (high-tech) devices and many sports are made accessible to blind and visually impaired people by a combination of modifications of the rules and the use of a sighted assistant.

18.4.4.1 Participation in Sporting Events

For most sports, blind and visually impaired people participate separately from sighted people and have their own clubs, competitions, tournaments and rules. However, this separate participation has been criticised and a number of blind and visually impaired people consider that everyone should participate together. There are also sports, such as showdown and torball, that have been created specially for blind and visually impaired people and have been taken up by sighted people. In many team games, blind players use audio information to make judgements about the position of the ball. It is therefore important that there is no competing external noise, for instance from cheering or booing crowds, which could interfere with them hearing the ball. This may lead to different atmospheres at sporting events for blind and sighted players.

Many countries have organisations of blind and visually impaired sportspeople, of which British Blind Sports is an example. These national organisations are frequently members of the International Blind Sport Federation (IBSA). The IBSA covers the following sports:

Athletics	Alpine skiing	Archery
Football	Goalball	Judo
Nine pin bowling	Nordic skiing	Powerlifting
Shooting	Showdown	Swimming
Tandem cycling	Ten-pin bowling	Torball

Participation in IBSA sporting activities is based on the following IBSA sight classifications:

- B1: 'totally blind', no light perception in either eye up to light perception, but the inability to recognise the shape of a hand at any distance or in any direction.
- B2: 'partially blind', from the ability to recognise the shape of the hand to a visual acuity of 2/60 or a visual field of less than 5° in the better eye after correction.
- B3: 'partially sighted', from visual acuity above 2/60 to a visual acuity of 6/60 or a visual field of less than 20° in the better eye after correction.

However, in the case of archery the classifications B1, B2 and B3 are based solely on visual acuity and there is an additional category of VI Open for people with visual field not greater than 20°.

As well as participating in sports, blind and visually impaired people are also sports fans and spectators. In the UK, the Visually Impaired Supporters Association was formed to represent the views of blind and visually impaired sports fans and to provide guidance on a range of issues include ticketing, seating, commentary and signage. Audio description (see Section 18.4.1.3) is now available at some sporting events.

18.4.4.2 Ballgames, Including Football, Cricket and Golf

Cricket is played by blind people in many different countries, as well as internationally. Unfortunately, the rules vary in different countries. Since the first Blind

Cricket World Cup took place in India in 1998, the international rules have largely been based on the Indian national rules, meaning that many blind cricketers have to deal with new rules and a slightly different approach to the game when going from the national to the international level. For instance, the UK game uses over-arm bowling, whereas the international game uses underarm bowling.

In the UK, a size three football rather than a cricket ball is used to make it more easily visible to players with some vision. The ball is filled with ball bearings to allow players to hear it and the wicket is larger than in the sighted game to make it both easier to see and touch for orientation. Additional audio cues are provided, such as the bowler asking whether the batsman is ready before beginning the run-up and shouting 'play' as the ball is released and the ball being required to bounce once and twice respectively when being bowled to a partially sighted and totally blind batsman. A blind fielder is allowed to catch the ball after it has bounced once and the number of runs scored by totally blind players is doubled.

The ball used in international cricket contains beads so players can hear it, but is the same size as a standard cricket ball. Allthough it is considerably less bouncy than the UK blind cricket ball, it is required to bounce on both halves of the wicket on its way to the batsman or batswoman. International cricket is played by two teams of 11 players, with at least four totally blind players (B1s), three partially blind players (B2s) and, at most, four partially sighted players (B3s). B1 batsmen should have a runner and B2 players have the option of having a runner.

The international version of blind football is called futsal. The game is played by two teams of five players, with four totally blind (B1) players and a goalkeeper who is either partially sighted (B2 or B3) or fully sighted and who may act as a guide. Team squads have a maximum of 13 members and consist of the players, a maximum of five substitutes, a coach, an assistant coach and a doctor or physiotherapist. The pitch is 18–22 m × 38–42 m and the longer sides are marked by kickboards of a height of between 1 and 1.2 m. This is considerably smaller than a standard football pitch, which should be 45–90 m × 90–120 m. The ball contains a sound system that does not impede its normal running, rolling and bouncing. In the interests of safety, this system is also required to make a noise when the ball is spinning on its own axis or spinning through the air.

A public address system is set up on top of the timekeeper's table and used to indicate that play has stopped and provide information on all incidents, including fouls, substitutions and time-outs. It is also used to ask the spectators to remain silent to enable players to hear the ball. Each team has a guide behind the opponent's goal, who guides the attacking players of the team, including by audibly indicating the position of the goalposts before a shot is taken with the ball stationary. The guide is not permitted to enter the field of play or protest the referees' decisions. The goalkeeper can also act as a guide when their team is in the third of the field closest to the goal being defended. At the referees' discretion, the goalkeeper may guide and orientate other players when free kicks, penalties or double penalties are being taken and organise the wall and position the players when a shot is taken at their goal. The goalkeeper is required to do this from the first third of the field of play.

There is also a version of football for blind people for B2/B3 players. In this case, it is important that reflections from sunlight or artificial light on the field of play are avoided and variations in light intensity on different parts of the field of play are prohibited. The ball can be any colour that facilitates locating it. The teams of five players have at least two B2 players. The goalkeeper can either be B2, B3 or fully sighted.

Blind golf is played all over the world and the first match took place in the 1920s. A sighted person, called the guide, describes the distance and direction of the ball and the characteristics of the hole, as well as helping the blind golfer align the club head behind the ball prior to making a stroke. The guide is not required to be a good golfer. The blind golfer then has to use this information to hit the ball in the right direction with an appropriate amount of force. Blind and visually impaired golfers are also allowed to ground their club in a hazard. Blind golf competitions are played in classes using the standard B1–B3 categories of visually impairment.

18.4.4.3 Swimming, Skiing, Athletics, Powerlifting and Tandem Cycling

Athletics is the most widespread IBSA sport and is practised in more than 70 countries. Competitions are based on the IBSA classification of B1, B2 and B3 athletes, with the standard rules of the International Association of Athletics Federations (IAAF) for B3 athletes and some modifications for B1 and B2 athletes. Most of the modifications relate to the provision of assistance, for instance by running guides, or callers providing auditory information.

Swimmers compete in the three IBSA sight categories, with B1 swimmers required to wear darkened goggles. The team coach directs the takeover at relays and allowances are made for B1 swimmers being too close to a lane line for technically correct arm strokes or touches in butterfly or breaststroke. Although the rules allow the use of an approved electronic device, sighted guides or tappers positioned at each end of the pool generally indicate the end of the pool for B1 and some B2 and B3 swimmers. They use a rod with a firm foam tip and synchronise their tap with the swimmer's stroke movement and momentum. Blind swimmers are not penalised if they accidentally surface in the wrong lane after a start or turn and are permitted to continue in this lane if it is unoccupied or to be given verbal instructions to return to the original lane by the tapper.

Alpine (downhill) skiing for blind and visually impaired people is a team sport, as a sighted person guides the blind and visually impaired skiers through the race course. It provides one of the rare opportunities for blind people to move freely at speed.

Powerlifting is one of the few sports where blind people do not require modifications, assistive devices or a sighted assistant and can compete identically to sighted people. Blind and visually impaired people can participate in cycling by using tandem bikes and sitting on the back seat, so the sighted person at the front steers. There are also racing tandems and a range of competitions over different distances. Tandem cycling can also be purely recreational. As well as steering, the sighted partner can describe the landscape and points of interest when touring.

18.4.4.4 Archery, Ten-Pin Bowling and Bowls

Archery is the most recent sport to be given official status in the IBSA. Blind and visually impaired archers may use tactile sighting systems and foot locators, as well as a sighted spotter. The spotter sets up and adjusts the sighting aid and foot markers before the start of the competition and quietly informs the archers of the strike of each arrow on the boss using the clock face method. The back of the hand tactile sighting aid is a pressure button type device placed on a stand in front of the archer. It gives a reference point for elevation of the bow by its contact with the back of the bow hand. The long rod system also provides information on elevation, but, instead of making contact with the bow hand, the long rod protruding from the bow riser is guided into a hook type device on the stand. Iris is an electronic sighting device developed in France in the early 1980s. It comprises a transmitter on the bottom of the target and a receiver on the bow, connected to a battery pack. The archer receives the transmitted signal over headphones. Increasing pitch indicates that the aim is approaching the centre of the target. Coaches considered that this was the best method, as it allowed visually impaired archers to be coached in a similar way to sighted archers with a sighting device on the bow. However, the Iris was very expensive and required considerable time to achieve an accurate aiming point, as well as involving numerous wires. Other than in France, it is no longer used, largely due to the high cost of the device. A laser device has also been developed, but high costs have prevented any significant usage. Foot markers are used to enable the archer to retain the same position when returning to the shooting line.

Ten-pin bowling is very popular among blind people and is played in more than twenty countries. Blind bowlers can use any bowling centre, but require either a guide rail or sighted guidance. When sighted assistance is used, the assistant aligns the bowler on the spot from which they wish to execute their deliveries. Guide rails are made of wood or lightweight tubular metal and are designed for easy assembly and disassembly. The rails are placed alongside the bowling approach, extending back from the foul line. The weight of the bowling balls holds the guide rails in place on the bowling approach and they can be used in any bowling centre, without damaging the lanes or interfering with the automatic bowling equipment. The bowler slides one hand along the smooth surface of the rail and delivers the ball with the other hand. A sighted assistant is generally required to inform the blind bowler which pins have been knocked down and which left standing by calling out the numbered locations of the pins. This information allows a blind bowler to know where to roll the next ball or to decide on how to modify the delivery of the next ball.

Bowling is played on a square bowling green surrounded by a shallow ditch. The game starts with one bowler placing the mat and rolling the jack to the other end of the green as a target. Bowlers then roll their bowls from the mat towards the jack and build up the 'head'. Bowling for blind and visually impaired people involves the minor modifications of the mat being situated with its front end at the fixed distance of six feet (1.8 m) from the ditch and a fine green string running under the centre of the mat and being fixed at both ends. This is used to help the bowler judge angles. A sighted assistant called a 'marker' stands behind the jack and provides information on the position of each bowl relative to the jack in terms of a 12-h clock. This allows blind bowlers to develop a mental picture of the position of the different bowls in the head.

18.4.4.5 Games Designed for Visually Impaired People: Showdown, Torball and Goalball

A number of ball games with an acoustic ball have been designed specifically for blind and visually impaired people, but can also be played by sighted people wearing a blindfold. Showdown was developed by Joe Lewis, a totally blind Canadian in the 1960s as a sport that could be played recreationally or competitively by blind people without assistance. It requires a specially designed table, two paddles, a special ball into which metal *bee bees* have been inserted and a room of suitable size to take the table. A glove is sometimes worn on the batting hand. The *bee bees* rolling round inside the ball indicate its location during play. The game involves batting the ball off the side wall, along the table, under the centre screen and into the opponent's goal. The first player to reach 11 points, leading by 2 or more points, wins. A player scores two points for a goal and one point when their opponent hits the ball into the screen or off the table or touches the ball with anything but the bat. Showdown is being played in many countries throughout the world and the IBSA is encouraging regional and national tournaments as part of a campaign for international recognition at the Paralympics.

Torball is a team sport for blind and visually impaired people developed in the 1970s. It is played by two teams of three players on a rectangular court 7 m × 16 m with a goal at either end. The ball only weighs 500 g and is pumped up with air, making play tricky and fast. It contains a bell, so players can detect its location, and must be thrown under three cords which are taut across the court. Each team tries to throw the ball across the opponent's goal, while the other team tries to prevent this from happening. The game is currently played by about 1200 people, in about 30 countries. It developed from Goalball, then known as Torball, which was based on Rollball, and invented in 1946 to help in the rehabilitation of blinded war veterans. This led to two varieties of Torball in central Europe. The older version was played with a 2-kg ball and later called Goalball, the English translation of the German name Torball, whereas the more recent version kept the name Torball and used a lighter ball of 500 g.

Goalball is also an IBSA sport and is also played by two teams of three players. The court is slightly larger than the torball court at 9 m × 18 m. It is marked out on the floor of a gymnasium and there is a goal at either end of it. The aim is to score goals by rolling the ball across the other team's goal line. The ball is heavier than the one used in torball at 1.25 kg. It contains noise bells and has eight holes. All players are required to wear eyeshades throughout the game, which lasts 20 min, divided into two halves of 10 min.

18.4.4.6 Accessible Sports Equipment

Accessible sports equipment can be divided into two main categories:

- Low-tech mechanical devices, including balls containing mechanical noisemakers and archery sighting aids. Many of these devices have been mentioned in the discussion of the different games and sports.
- Higher tech talking devices, including step counters and compasses.

18.4.4.6.1 Audible Balls

These include the following products supplied by the RNIB:

- An inflatable white rubber football, containing carbon steel ball bearings, which rattle when the football is shaken or moved. The ball has a valve and can be reinflated with a bicycle pump.
- A blue rubber ball for goalball (See Section 18.4.4.5) with eight holes and containing a metal bell for location in play.
- An orange inflatable rubber ball containing bells that can be used in a gym or a swimming pool.

18.4.4.6.2 Talking Step Counter

Cobolt Systems distribute two talking pedometers. They can both be clipped to a belt, carried in a pocket or just carried. Both tell the user the distance walked or run. One version has FM radio and also announces the calories consumed and the elapsed time. The announcements can be heard at preset intervals, or at the touch of a button. The other version also provides a talking alarm with a choice of four sounds and a selectable music function with its tempo dependent on how fast the user is walking. However, it may not always be very accurate, as it requires the user to enter their average step length. Utmost Technology Corp also distribute a talking pedometer, with a voice announcement and display of the number of steps and the distance travelled. It can measure up to 99,999 steps and 1999.99 km or miles. Seven melodies can be played while walking or jogging, with the music tempo synchronised with the walking or jogging speed. The pedometer also has a talking alarm clock with four different alarm sounds.

Cobolt Scientific's talking digital skipping rope has a calorie counter, jump counter and timer. The voice announces the number of calories consumed, the number of jumps, the workout time and the time remaining until the end of the workout, with a maximum setting for the workout time of 99 min. It will speak at preset intervals or on pressing the talk button. There are also four display modes for the number of jumps, number of calories, elapsed time and countdown timer. The rope length can be adjusted for the user.

18.4.4.6.3 Magnetic Compass and Devices for Visually Impaired Sailors

There are several magnetic compasses with speech or other audible output and a range of different features. Most of them are eight point compasses and there are also versions designed to be used by visually impaired sailors. A different approach is provided by the Feelspace tactile compass from WiFi-ArT.com, which consists of a wearable belt that enables the user to feel their orientation in space by vibro-tactile vibrations. The electronic compass controls the vibrators in the belt. The element pointing north is always vibrating slightly to give the user constant information about their position relative to north. Another tactile compass is the Lensatic tactile compass from Assistech, which has three tactile dots to mark north and a cover for protection.

Talking compasses for stationary use include the Robotron Columbus talking compass, generally referred to as the C2 (see Figure 18.8). It is a miniature hand-held battery operated eight-point compass. It has digitised speech output with a choice of two languages or voices that can be chosen at the time of the order, with defaults of English and Spanish. The device has a semi transparent plastic case to allow the use of a series of colour LEDs (light emitting diodes) by people with low vision. There is a carrying strap which can be looped round the user's wrist to prevent them losing it. The C2 should be held as level as possible to avoid errors, but there is no audible out of level or low battery indicator. There are two controls. A small three-position slide switch on the left side of the unit turns it on and off and determines the choice of language. A domed button on the top of the unit gives speech output of the compass point, which is repeated as long as the button is held down.

Talking compasses for visually impaired sailors include the Talking Compass Repeater from Tinley Marine Electronics (see Figure 18.9), who also supply the Talking Depth Repeater for giving depth measurements. The Talking Compass Repeater has an internal loudspeaker with a clearly annotated set of push button

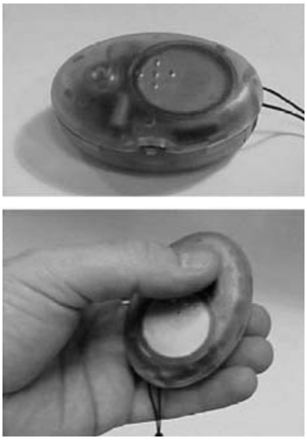


Figure 18.8. C2 Talking Compass (photograph reproduced by kind permission of Robotron Group, Australia)

Figure 18.9. Talking Compass (photograph reproduced by kind permission of Tinley Electronics Ltd. UK)

controls that set the repeat rate and the volume level of the audio output. There is also an optional external speaker cable for additional audio output of the compass readings.

OnTrack is software which can be used with any standard GPS (global positioning system) connected to a computer to convert course heading information into audible tones giving the user the direction and how much to turn to follow the desired course. It will run on Windows 95 and above and palm and PCs.

The Marine Instrument Talker from Real Time Control can add speech capability to any instrument with National Marine Electronics Association (NMEA) 0183 output. This is a standard that defines electrical signal requirements, the data transmission protocol and time and specific sentence formats for a 4800 baud serial data bus. Applications of the Talker include talking depth sounders, talking boat speed indicators, talking compasses and talking wind direction indicators. The system repeat interval can be adjusted from continuous speech to a 4-min delay before repetition. The system can be set to speak the data whenever it changes, for instance in depth sounding. Various different formats can be set for speaking data and the selections and settings are retained when the Talker is switched off. Installation simply requires a speaker, the vessel's 12-V power supply and the NMEA output from the instrument to be connected.

18.4.5 DIY, Art and Craft Activities

Do it yourself (DIY) activities include painting and decorating, plumbing, putting up shelves, (re)wiring the house and mending broken tools and appliances. They are popular due to the costs of involving professionals and the sense of achievement if the process is successful. With sensible safety precautions in place, there is no reason why blind and visually impaired people should not engage in many aspects of DIY, however, the range of tools with audio output or tactile markings is still very limited. Some of these tools were discussed in Section 18.3.4 and measuring tapes are presented at the end of this section. The rulers, measuring sticks and callipers presented in Section 18.2.2.5 may also be of interest. Some ordinary tools such as screw drivers, spanners and adjustable spanners can, in principle, be used without adaptation by blind and visually impaired people, though raised size markings would be helpful and are found on some spanners. Blind and visually impaired people may require clamps to hold the objects they are working on in place and make them easier to find. However, it would be useful to have a tactile device for measuring the size of screws, nuts, bolts and spanners. Similar principles to those of the knitting needle gauge considered below could be applied for screws and bolts, but a different design would be required for nuts and spanners.

Safety is a particular concern in many aspects of DIY, for instance those that involve high ladders, the use of heat, as in a blowtorch or a soldering iron, or involve electricity, as in electrical wiring tasks. Thus, there is a need for quidelines on safety procedures and access to expert advice before attempting many DIY tasks. There is also a role for recommendations as to what DIY activities visually impaired and blind people should think very carefully about undertaking. The availability of additional safety information and guidelines would be of equal benefit to sighted, as well as visually impaired and blind DIY enthusiasts. Although an outdoor activity and not normally classified as 'DIY', gardening seems to fit more logically in this category than in the sports and outdoor activities category. Therefore the various techniques and the few adapted tools for making gardening accessible to blind and visually impaired people are discussed in this section.

The term art and crafts covers a variety of activities, including painting, photography, sewing and knitting. The section includes consideration of a number of mechanical aids, some of which can be home made, as well as techniques and practical suggestions for sewing and knitting. There is also discussion of a range of tape measures, painting and photography.

18.4.5.1 Sewing and Knitting

Most of the adaptive devices available in this area are of the low technology type. For instance, there are several different designs of mechanical needle threaders, as well as so-called 'self-threading' needles. Self-threading needles have two holes, the upper one of which is open. The needle is placed in a vertical position in a pincushion, cork or even a bar of soap, the thread is wrapped round the two index fingers and the thumbs are used to locate the position of the needle and push the thread down onto the needle. Unfortunately, the open upper hole means that the needle unthreads very easily. Dental floss threaders provide a more robust and less bendable, though still flexible, alternative to wire needle threaders. Some sewing patterns are available in large print and/or Braille. Suggestions for creating or adapting existing patterns include using heavy brown paper, outlining the cutting markings with Elmer's glue and using masking tape to identify most of the other markings, with staples used to identify the location of notches and darts. A coloured broad tipped felt pen can be used to trace all cutting lines and markings to make them more visible (Woodward 1998).

There are a number of mechanical aids that can be used to enable blind and visually impaired people to sew straight and safely. These include a needle finger guard, which is a small three-sided metal bar attached to the sewing machine. It is pulled down in front of the needle when sewing and pushed up to the left of the needle when threading. A metal screw-on or magnetic seam guide can be attached to most sewing machines to guide sewing, but the screw-on type is often preferred as it is less likely to change position. A longer edge to guide sewing can be created by the combination of a metal seam guide with a pinhead guide made by placing straight pins horizontally onto the sticky side of masking tape with their heads extending over one side (Woodward 1998). Alternatively, a guide made from card can be used or a suitable ruler, such as the American Printing House brass ruler, glued to the side of the sewing machine.

Other suggestions for sewing and needlework include the following:

- Using large-headed glass pins and coloured knitting needles in contrasting colours to the working fabric or knitting fibre.
- Placing a tactile, brightly coloured magnetic tape along the seam marker allowance of a sewing machine.
- Using a basket or tray or an apron with several pockets to assemble scissors, pins, needles and spools, and using magnets to pick up any dropped pins or needles.
- Putting tissue paper between the fabric and a same colour button to show contrast.
- Storing reels of sewing thread in a multi-drawer storage box from a DIY store. Labels that could be used for marking the drawers are discussed in Section 17.2.1.

Enlarged knitting patterns can be produced on a photocopier and this will also remove the glare from glossy paper. There are also some knitting or crochet patterns available in large print, Braille and on tape. The RNIB produces a plastic knitting needle gauge, with circular holes cut to the circumference of 19 different sizes of knitting needle. The needle size is determined by finding the hole the needle fits into and then reading the Braille and large print markings of the imperial and metric sizes at the side of the hole. The Brynolf pocket counter from the RNIB can be used as a knitting counter (or to add sums of money up to £99.99). It is a black plastic manual counter with four separate columns containing white sliding counting discs and has a cord wrist strap. There is an international email list for blind knitters and crocheters.

18.4.5.1.1 Tape Measures

Cobolt Systems produce a 5-m (16-ft) tape measure with a resolution of 1 mm (1/16 in). The measurements are spoken automatically in English. Measurements can be added in memory up to a total of 99.99 m. The memory is retained when the unit is switched off. The tape can be set to zero at any position. Readings can be made in metres, centimetres, millimetres or feet and inches and converted between these units.

The Delux measuring tape from the Braille Superstore is a 5-ft (1.5-m) measuring tape with each 1 in (2.5 cm) marked by one round eyelet and each 1 ft (30.5 cm) marked by two eyelets (see Figure 18.10).

Figure 18.10. Tactile measuring tapes (photograph reproduced by kind permission of The Braille Superstore, Canada)

18.4.5.2 Painting and Photography

Visually impaired artists may use some of the low vision aids described in Chapter 12 to enable them to focus on part of a scene. A mount, consisting of cardboard with a cut-out hole the size of the painting in the centre, can then be used to view the whole picture to determine its balance. Each visually impaired or blind artist will work out their own style and technique to make best use of the available light, optic aids and residual vision and support and training can be very helpful in doing this.

The Seeing with Photography Collective is a New York-based group of photographers, who are visually impaired, blind and sighted. They use flashlights, darkness and a technique called painting with light. The blind photographer directs a sighted assistant to focus and compose the view frame. The camera shutter is kept open for several minutes in a darkened room while the sitter is slowly painted with a flashlight. This gives rise to a very distinctive photograph, characterised by luminous distortions and blurred and glowing forms.

The Talking Camera or Kurzweil-NFB Reader consists of a digital camera packaged with a pocket computer. It is designed to read documents and the camera has a field of view report feature which can be used to align the document and reader. The device is rather expensive to be used purely as a camera and further developments would be required to recognise people and other objects than documents and align them with the viewer.

18.4.5.3 Gardening

While few specific tools or assistive technology for blind gardeners seem to have been developed, a number of organisations for blind people provide gardening tips. One of the few special tools, devised by a blind gardener in 1968, consists of a right angle, marked with studs for measuring planting distances. It can be made of plastic piping or wood and used for helping to square corners, planting shrubs and marking an area for lawn maintenance, which generally requires two right angles to be combined, as well as measuring planting distances. In this case two right angles are put at opposite ends of a flower bed and joined with cane, string or another piece of wood or plastic.

The Kent Association for the Blind and Vision Magazine in the UK have provided the following useful suggestions:

- Methods to make it easier and safer to get around the garden, including using good illumination and coloured furniture, cutting back hedges and painting the edges of steps and handrails in contrasting colours. Shrubs, trees, scented plants, furniture, rustling plants, running water and wind chimes can be used as landmarks, but need to be easy to maintain, not get in the way and easy to distinguish from each other. However, the use of running water and wind chimes as landmarks could be irritating to neighbours.
- Safety suggestions, such as wearing protective goggles and gloves, having circuit breakers at the mains on all power tools, using lawn mowers with plastic blades and using link stakes rather than bamboo canes or covering the pointed tops of canes with, for instance, a pad made from discarded textiles.
- Marking tools with bright colours to keep track of them and carrying small tools in a white bucket.
- Using the middle section of plastic drinks containers as a protective collar round plants to distinguish between weeds and plants and protect plants while hoeing.
- Labelling plants with pegs with differently shaped tops, raised numbers or vegetable shapes made from plastic containers.
- Learning the shape of shrubs by feeling from the base upwards or by finding pruning points, while remembering that thorns point downwards.
- Marking fences, tool handles, posts and other items with white paint.

Although not specially designed for blind and visually impaired people, the automatic robotic lawnmower from Grassland Services may be of interest. A thin wire needs to be pegged down round the perimeter of the lawn, but gradually disappears into the lawn. The battery-powered rotary mower is heavy, but can be driven with a remote control onto the lawn. Once on the lawn, pressing the green button starts it automatically cutting the lawn and mulching the cuttings. The blades are height adjustable, but the mower will not work on slopes over 15° and is best suited to well kept lawns. The mower is currently expensive, but the price may reduce, as it is proving popular.

The RNIB produces a booklet called Gardening without Sight in large print, Braille and tape versions. There are also schemes for pairing up sighted and visually impaired gardeners, and road shows and residential courses for visually impaired gardeners. In the UK, there is a National Blind Gardeners Club.

18.5 Chapter Summary

This chapter has used the activities module of the CAT model to classify the accessibility devices and solutions for overcoming the barriers encountered by visually impaired and blind people in carrying out educational, employment and recreational activities. As well as providing a useful structure for the chapter, this approach has provided an illustration of the value of the CAT model in discussing and describing assistive technology provision. The employment and education aspects of the employment and education category have been separated, giving the chapter three main sections on education, employment and recreational activities.

An overview of the chapter shows both the great ingenuity of blind and visually impaired people in developing their own techniques and accessibility solutions and the relatively few systems developed using design for all, specifically for blind and visually impaired people or made accessible to them through modifications or assistive devices.

In the area of education, the focus has generally been on blind people as students rather than as teachers, though the latter is equally important. This area is marked by both sophisticated high-tech solutions, such as talking graphical and Braille calculators and computer-based talking or Braille systems for representing mathematics and low-tech mechanical devices, such as tactile rulers and metre sticks and Moon number sheets.

The further divisions of the employment category into professional and person-centred, scientific and technical, administrative and secretarial, skilled and unskilled (manual) trades, and outdoor work formed the basis of the subsections of Section 18.3. Information about employment practices, assistive technologies and ingenuous techniques and adaptations by blind and visually impaired people in the different areas is very uneven. For many employment activities, it is access to information technology and print that is particularly important and this is discussed in Chapters 12, 13 and 15. This is the case for professional and person-centred activities and, to a large extent for administrative and secretarial activities, though the section does discuss a device developed by the Kentucky Department for the Blind for accessing LCDs on equipment, such as fax machines. The section on scientific and technical activities focuses on making laboratories accessible. A range of accessible, adapted and assistive equipment, devices and machinery was presented in the section on skilled and unskilled (manual) trades. This included a battery tester as well as optical and digital magnification of equipment and machinery status indicators.

Although recreational activities are very important for health and wellbeing, they tend to have a low profile in discussion and development work on assistive technology. Therefore, though they are the last activity category described in the book, it is very much a case of last, but not least. The CAT model was used to divide the recreational activities section into the following subsections: accessing the visual, audio and performing arts; games, puzzles, toys and collecting; holidays and visits; sports and outdoor activities; and DIY and craft activities. The category of accessing the visual, audio and performing arts is generally marked by high-tech solutions, such as talking DAB radios, though there are also intermediate tech solutions, such as optical television screen magnification. One of the most important solutions in this area is audio description, which illustrates the felicitous combination of human assistance to describe the programme or event and technology to transmit the description.

A number of self-voicing computer games have been developed, but low tech solutions, involving Braille markings, tactile dice and the use of shape and colour to distinguish pieces and other objects, are more common in the area of games, puzzles, toys and collecting. There are a number of tactile tours and some accessible information is being provided for blind and visually impaired visitors to museums, galleries and heritage sites, but the potential for making these sites interesting as well as fully accessible to blind and visually impaired people has not yet been fully realised. The participation of blind and visually impaired people in sport is generally supported by sighted assistance combined with low-tech devices such as noise making balls and slight modifications to the rules and/or equipment. However, high-tech devices, such as talking compasses and depth measurement devices, have been developed for outdoor activities such as sailing. There are also a number of games that have been developed specifically for visually impaired people, such as torball and showdown.

Many of the solutions for DIY activities were discussed in the section on skilled and non-skilled (manual) trades. Craft areas, such as knitting and sewing, use mechanical solutions, such as knitting needle gauges and tape measures with eyelet markings, some of which can be made at home. Gardening is marked by much ingenuity and practical suggestions arising from experience. Although not designed specifically for blind and visually impaired people, an automatic robotic lawnmower (the mower described was from Grassland Services) may be of interest.

Questions

Q.1 Describe how the CAT model can be used to provide a framework for the activity categories of education and employment and recreational activities.
Q.2 Discuss how the operation of office equipment can be made accessible for blind people?
Q.3 What is meant by the term 'audio description'? What are its main applications? Discuss briefly the main principles of audio description.
Q.4 Compile a list of the main methods used to make five different sports accessible to blind and visually impaired players.
Q.5 Discuss the technologies and techniques used by blind and visually impaired people in DIY and craft activities.

Projects

P.1 Consider the design of a DIY workshop to be accessible, easily usable and safe for blind people. Draw up guidelines for accessibility, usability and safety which cover the following areas:
   (a) Layout
   (b) Decor
   (c) Equipment
   (d) Cleaning
   (e) Utility services

P.2 Discuss and evaluate the various approaches to making mathematics and sciences accessible to blind and visually impaired people.

P.3 Choose an activity area where there are few existing solutions or where you consider that existing solutions are unsatisfactory and there is a role for new electronic assistive technology:
   (a) Draw up a survey questionnaire or other methodology for consulting end user groups on their requirements and obtain appropriate ethical approval for using it.
   (b) Apply the questionnaire or other methodology to obtain feedback from end users.
   (c) Apply this feedback to draw up design specifications for the new assistive device.
   (d) Construct the device
   (e) Test and modify the device. This should involve a number of stages, ending with end-user testing, but this should only be carried out after ensuring that the device is safe and obtaining authorisation from the appropriate ethics committee(s).You also need to ensure that testing does not expose end-users to any risk, as well as paying appropriate attention to data confidentiality, providing full information and consent issues.

P.4 Consider a blind person who is moving into higher education and hopes to have an interesting and well paid career with good promotion prospects at the end of it. Use the CAT model to investigate some of the issues raised by different education and career choices.

References

Alonso, F., Fuertes, J.L., Gonzáles, A.L., and Martínez, L.A., 2006, SBT: a translator from Spanish mathematical Braille to MathML, Proc.s 10th Int. Conf, ICCHP 2006, Computers Helping People with Special Needs, pp. 1122–1130, Springer

Annamalai, A., Gopal, D., Gupta, G., Guo, H., and Karshmer, A., 2003, INSIGHT: a comprehensive system for converting Braille based mathematical documents to Latex, In: Universal Access to HCI – Inclusive Design in the Information Society, C. Stephanidis (ed.), Vol. 4, pp. 1245–1249, Mahwah, NJ

Anon, undated, Bureau of the Census: Survey of Income and Program Participation, 1994–95, http://www.census.gov/hhes/www/disable/dissipp.html

Archambault, D., and Moço, V., 2006, Canonical MathML to simplify conversion of MathML, Proc. 10th Int. Conf, ICCHP 2006, Computers Helping People with Special Needs, pp. 1122–1130, Springer

Archambault, D., Fitzpatrick, D., Kupta, G., Karshmer, A.I., Miesenberger, K., and Pontelli, E., 2004, Towards a universal maths conversion library, Proc. ICCHP 2004, Paris, France, Lecture Notes in Computer Science, Vol. 3118, pp. 664–669

Barker, A., 2006, PhD thesis, Dept. Electronics and Electrical Engineering, University of Glasgow, Glasgow, UK

Batusic, M., Miesenberger, K., and Stoeger, B., 1998, Labradoor, a contribution to making mathematics accessible for the blind, Proc. ICCHP 98, pp. 307–315, Oldenbourg

CML, undated, http://inova.snv.jussieu.fr/umcl-demo

Crombie, D., Lenoir, R., McKenzie, N., and Barker, A., 2004, Math2braille: opening access to mathematics, Proc. ICCHP 2004, Paris, France, Lecture Notes in Computer Science, Vol. 3118, pp. 670–671

EBU, 2001, Survey on the employment of blind and partially sighted people in Europe (2001), European Blind Union, http://www.euroblind.org/fichiersGB/surveymb.htm

Ferres, L., Parush, A., Roberts, S., and Lindgaard, G., 2006, Helping people with visual impairments gain access to graphical information through natural language: the iGraph system, Proc. 10th Int. Conf, ICCHP 2006, Computers Helping People with Special Needs, pp. 1122–1130, Springer

Fitzpatrick D., 2006, Mathematics: how and what to speak, Proc 10th Int. Conf, ICCHP 2006, Computers Helping People with Special Needs, pp. 1199–1206, Springer

French S., 1995, Visually impaired physiotherapists: their struggle for acceptance and survival, Disability and Society, Vol. 10, No. 1, pp. 3–20

Gardner, J.A., Ungier, L., and Boyer, J.I., 2006, Braille math made easy with the Tiger formatter, Proc. 10th Int. Conf, ICCHP 2006, Computers Helping People with Special Needs, pp. 1215–1222, Springer

Gougoux F., Lepore, F., Lassonde, M., Voss, P., Zatorre, R.J., and Belin, P. 2004, Neuropsychology: pitch discrimination in the early blind, Nature Vol. 430, pp. 309–310

Hersh, M.A., Baker, N., Macleod, M., and Weightman, B. 2004, The accessible electronics laboratory, Int. Conf. on Innovation, Good Practice and Research in Engineering Education, Wolverhampton, England, A. Halstead *et al.* (eds.) Vol. 2, pp. 175–181

Hetherington, K., 2003, Accountability and disposal: visual impairment and the museum, Musem and Society, Vol. 1, No. 2, pp. 104–115

Hirst, M., and Baldwin, S., 1994, Unequal Opportunities, Growing up Disabled, SPRU Editorial Group, Science and Technology Policy Research, University of Sussex

ITC, 2000, ITC guidance on standards for audio description, Independent Television Commission

Jansson, G., Gergamasco, M., and Frisoli, A., 2003, A new option for the visually impaired to experience 3D art at museums: manual exploration of virtual copies, Visual Impairment Research, Vol. 5, No. 1, pp. 1–12

Keil. S., and Clunies-Ross, L., 2003, Survey of education provision for blind and partially sighted children in England, Scotland and Wales in 2002, RNIB, London

Kirchner, C., Schmeidler, E., and Todorov, A., 1999, Looking at Employment Through a Lifespan Telescope: Age, Health and Employment Status of People with Serious Visual Impairment, Mississippi State, MS: Rehabilitation Research and Training Center on Blindness and Low Vision

LBA, undated, http://www.lambda.project.org

Moço, V., and Archambault, D., 2003, Automatic translator for mathematical Braille, In: Universal Access to HCI – Inclusive Design in the Information Society, C. Stephanidis (ed.), Vol. 4, , pp. 1335–1339, Mahwah, NJ

Petré, L., Sinclair, D., Hopkins, A., Couthard, L., Walker, E., Stallard, G., Morris, M., Jones, G., Hanson K., Staines R., Herbert T., Porter V., Myers R., Klein G., and Hobbs, B., 2006, Digital TV equipment: vulnerable consumer requirements, A report by the Consumer Expert group to government and Digital UK

RNIB, 2006, Information from RNIB National Centre for Tactile Diagrams, http://www.nctd.org.uk/MakingTG/index.asp

Schweikhardt, W., Bernareggi, C., Jessel, N., Encelle, B., and Gut, M., 2006, LAMBDA: a European system to access mathematics with Braille and audio synthesis, Proc. 10th Int. Conf, ICCHP 2006, Computers Helping People with Special Needs, pp. 1223–1230, Springer

Smith, A., and Twomey, B., 2002, Labour market experience of people with disabilities, Labour Market Trends, Vol. 110, No. 8, pp. 415–428

Smith, T., 2002, Diversity and disability: exploring the experiences of vision impaired people, Equal Opportunities International, Vol. 21, No. 8, pp. 59–72

Tate, R., Smeeth, L., Evans, J., and Fletcher, A., 2005, The prevalence of visual impairment in the UK: A review of the literature. Report commissioned by RNIB, http://www.rnib.org.uk/xpedio/groups/public/documents/PublicWebsite/public_prevalencereport.doc

Woodward, C. 1998, Clothing care and sewing techniques for visually impaired or totally blind students, Texas School for the Blind and Visually Impaired http://www.tsbvi.edu/Education/clothing.htm

Yoshimoto, T., 1901, Massage by the blind in Japan, The Blind, Vol. 16, pp. 292–293

Resources

RNIB publications on leisure, DIY, craft, sport and holiday activities for visually impaired and blind people

Leisure ideas for people with a sight problem
Do it for yourself – an introduction to DIY for people with sight problems
Painting from a new perspective
Getting on with gardening
Leisure in Focus
Creative writing and access to literature
Discovering sports
Knitting information pack

Sources of funding for leisure-based projects

Going on Holiday
Holiday Ideas
Helping you to take a holiday – How to find funding and escorts for a holiday
Museums and galleries – a user guide
Music appreciation for people with sight problems

RNIB publications for professionals

Activate – Helping people with sight problems to access gyms
Designing gardens and nature trails
Leisure in Focus
Welcoming your visually impaired customers, tourist industry pack
Talking Images Guide – museums, galleries and heritage sites: improving access for blind and partially sighted people
Talking Images Research – museums, galleries and heritage sites: improving access for blind and partially sighted people
Digital Audio Broadcasting development report 2000

Products and devices

ABLEDATA, product database, USA, http://www.abledata.com
American Foundation for the Blind, USA, http://www.afb.org/
American Printing House for the Blind, USA, http://www.aph.org
American Thermoform Corporation, USA, http://www.americanthermoform.com
Ann Morris Enterprises, USA, http://www.annmorris.com/
Assistech, USA, http://azhearing.com/
Braille Superstore, Canada, http://www.braillebookstore.com Braille and talking products
Brytech, Canada, http://www.brytech.com/ Banknote readers and health monitoring products
CareTec Gmbh, Austria, http://www.caretec.at
Cobolt Systems Ltd., The Old Mill House, Mill Road, Reedham, Norwich, Norfolk, NR13 3TL, U.K. Website: http://www.cobolt.co.uk
Compusult Ltd, 40 Bannister Street, P.O. Box 1000, Mount Pearl, NL, Canada A1N 1W1 Website: http://www.Compusult.net
Duxbury, http://www.duxburysystems.com
Exceptional Teaching Aids Inc, USA, http://www.exceptionalteaching.com
Games for the Blind, USA, http://www.gamesfortheblind.com/
Handy Tech Eletronik GmbH, Germany, http://www.handytech.de/en/
Institut de Réadaption en Déficience Physique de Québec, Canada, http://www.irdpq.qc.ca/

IntelliTools Inc, USA, http://www.intellitools.com
Kentucky Department for the Blind, USA, http://blind.ky.gov/
Lighthouse International, USA, http://www.lighthouse.org
National Federation of the Blind, USA, http://www.nfb.org
Nebula Electronics, UK, http://www.nebula-electronics.com/
Modern World Data Ltd, UK, http://www.modern-world-data.com/
Omega Engineering, USA Inc, http://www.omega.com
Portset, UK, http://www.portset.co.uk/
Real Time Control, Australia, http://home.alphalink.com.au/~derekw/realtime.htm
Royal National Institute for the Blind (RNIB) (online shop), http://onlineshop.rnib.org.uk/
Robotron Sensory Tools Divison, Australia, http://www.sensorytools.com/
RYA Sailability Equipment Directory, UK, http://www.sailability.co.uk/
Seiko, UK, http://www.seiko.co.uk
Spektra v.d.n., Czech Republic, http://www.spektravox.cz
Tacktick, UK and France, http://www.tacktick.com/
TechReady, UK, http://www.techready.co.uk
Tinley Marine Electronics, UK, http://www.tinleyelectronics.com
Ultmost Technology Corporation, Republic of China, http://www.asia.globalsources.com/ultmost.co
Viewplus, USA, http://www.viewplus.com/
Zychem Ltd, UK, http://www.zychem-ltd.co.uk/

Audio description

Audio Description International, http://www.adinternational.org
Vocal Eyes, UK, http://www.vocaleyes.co.uk

Sports associations

Blind Sailing International, http://www.blindsailing.org
British Blind Sports, UK, http://www.britishblindsport.org.uk
International Blind Sport Federation, http://www.ibsa.es/eng/
NFB Sports and Recreation Division, USA
http://www.nfb.org/nfb/Divisions_and_Committees.asp

Professional and trade associations

Association of Blind and Partially Sighted Teachers and Students (ABAPSTAS), UK http://www.abapstas.org.uk
Association of Blind Piano Tuners (ABPT), UK, http://www.uk-piano.org/abpt
NFB Agriculture and Equestrian Division, USA, http://www.nfb.org/nfb/Divisions_and_Committees.asp
Blind Musicians, http://www.nfb.org/nfb/Divisions_and_Committees.asp
Blind Professional Journalists Group, http://www.nfb.org/nfb/Divisions_and_Committees.asp
NFB Human Services Division, USA, http://www.nfb.org/nfb/Divisions_and_Committees.asp, Membership organisation for blind people working in counselling, personnel, psychology, social work, psychiatry, rehabilition and other social science and human resources fields
National Association of Blind Enterpreneurs, http://www.nfb.org/nfb/Divisions_and_Committees.asp, Membership organisation of blind people who are self-employed, directing businesses or interested in starting a business
National Association of Blind Lawyers, USA, http://www.nfb.org/nfb/Divisions_and_Committees.asp, Membership organisation of blind lawyers, law students, judges and others in the law field
National Association of Blind Merchants, http://www.blindmerchants.org/, Membership organisation of blind people working either or a self-edmployed basis or in the Randolph-Sheppard vending program

National Association of Blind Office Professionals,
http://www.nfb.org/nfb/Divisions_and_Committees.asp,
Membership organisation of blind secretaries and transcribers, including medical and paralegal transcription, office workers and customer service personnel
National Association of Blind Piano Tuners, http://www.nfb.org/nfb/Divisions_and_Committees.asp, Membership organisation of blind people working as professional piano tuners or interested in this career
National Associations of Blind Rehabilitation Professionals, USA,
http://www.nfb.org/nfb/Divisions_and_Committees.asp
National Association of Blind Students, http://www.nfbstudents.org, Provides support and information to blind college and university students
National Federation of the Blind in Computer Science,
http://www.nfb.org/nfb/Divisions_and_Committees.asp,
National organisation of blind people knowledgeable about computer science and technology
National Organization of Blind Educators, http://www.nfb.org/nfb/Divisions_and_Committees.asp, Membership organisation of blind teachers, professors and instructors at all levels of education
Public Employees Division, http://www.nfb.org/nfb/Divisions_and_Committees.asp, Blind people working in the public sector
Science and Engineering Division, http://www.nfb.org/nfb/Divisions_and_Committees.asp, Blind people with expertise and experience in science and/or engineering. The organisation provides support for them to carry out careers in these areas

Mathematics and graphics

Lambda Project, http://www.lambdaproject.org/ASP/index.aspx?IDMenuAPP=0
RNIB National Centre for Tactile Diagrams, UK, http://www.nctd.org.uk

Gardening and tourism

Kent Association for the Blind, UK http://www.kab.org.uk/
National Blind Gardeners Club, UK http://www.thrive.org.uk/specific-services-blind-club.asp
NFB Travel and Tourism Division, http://www.nfb.org/nfb/Divisions_and_Committees.asp

Blind writers and theatre companies

Extant, Arts Production Company, UK, http://www.extant.org.uk/
NFB Writers Division, USA, http://www.nfb.org/nfb/Divisions_and_Committees.asp
Theatre by the Blind, USA, http://www.tbtb.org/
Theatrical Company of the Blind and Visually Impaired, Croatia, http://www.hr/darko/etf/blind.html

Biographical Sketches of the Contributors

Elizabeth M. Ball

Liz Ball began a PhD in December 2002 at the Ergonomics and Safety Research Institute, Loughborough University. Since April 2004 Liz has also worked part-time as a project co-ordinator in the Campaigns Team at Sense. Before this, in 2002 Liz completed an MSc in Information Technology (Human Factors) and in 1999 graduated with a BSc in Psychology, also from Loughborough University. In 1998 she was awarded the departmental prize for psychology and in 1999 was awarded the Sir Robert Martyn Faculty Prize for academic achievement and service to the community. Liz has also worked as an information officer for Rethink and as a learning support assistant in a mainstream school.

Liz's research, for both her MSc and PhD, is in the field of orientation and mobility for blind and partially sighted people. Her MSc project looked at user requirements for navigation systems to be used by visually impaired pedestrians. Her PhD is looking at blind and partially sighted people's experiences of learning to become independent travellers and is focussing on orientation and mobility training.

Paul Blenkhorn

Paul Blenkhorn is currently Professor of Assistive Technology in the School of Informatics, University of Manchester (formerly the Department of Computation, UMIST) where he has been for the past 12 years. Prior to that he was first a Research Fellow at the Research Centre for the Education of the Visually Handicapped (1983–1986) and then Director of R&D and one of the founders of the assistive technology company Dolphin Systems (1986–1991).

For many years his interests have been in the two areas of computer access and the use of technology to support people with profound and multiple disabilities. He has developed screen readers and magnifiers on several platforms including the BBC micro, the NEC 8201 and MS DOS (in the 1980s) and more recently for Microsoft Windows.

Nicholas A. Bradley

Since August 2006, Nick has been working as a Research Fellow for NHS Education for Scotland based in Glasgow, Scotland. He undertakes research into the development of techniques for improving the quality of patient care across various health care professions. Prior to this, Nick managed a usability team at a contact centre software company called Graham Technology headquartered in Inchinnan, Scotland. In 2005 he completed a PhD in the Department of Computer and Information Sciences at Strathclyde University. His research involved investigating Human Computer Interaction (HCI) issues of context-aware mobile systems, particularly those used by people with visual impairments to navigate and orientate. Prior to this, Nick graduated in 2001 with a BSc (Hons) in Ergonomics at the Human Sciences Department at Loughborough University. He also received a Diploma of Professional Studies (DPS) after completing a student placement at Nickleby HFE (Human Factors and Ergonomics) Ltd, Glasgow, UK. On his graduation day he was awarded the Departmental Prize for Ergonomics. He is a Graduate Member of the Ergonomics Society, a research affiliate of the Transport Research Institute and a member of the Glasgow Interactive Systems Group (GIST).

David Crombie

David Crombie has an MSc in Economics from University of London and works in the Utrecht School of the Arts, Research Institute on Digital Cultures (URIDC) as well as running the International Projects department at DEDICON (formerly FNB Nederland). He is a founding member of the Openfocus Institute in Amsterdam and chairs the European Accessible Information Network. He has worked on many European accessible content projects and has collaborated with many of the major European organisations and individuals in this field. The HARMONICA project aimed to provide a solid strategic framework for networked access to music and related multimedia services he co-ordinated the MIRACLE project to build a worldwide on-line library of music in alternative formats. He is responsible for the accessibility components of the WEDELMUSIC, TEDUB, MULTIREADER and MUSIC NETWORK projects. He is currently involved in several cross-sectoral domains and initiatives relating to accessible software architectures; information modelling; and intelligent knowledge management strategies.

Mark D. Dunlop

Dr Dunlop is currently a Senior Lecturer in the Department of Computer and Information Sciences, Strathclyde University, Glasgow. Before joining Strathclyde, Mark was a senior scientist at the Danish Centre for Human Machine Interaction at the Risø National Laboratory. Prior to Risø, he was a lecturer in Computing Science at the University of Glasgow. He has co-organised three meetings of the International Conference Series on MobileHCI (MobileHCI 99, 01 and 04), has chaired the MobileHCI steering committee and is currently a member of the board of the journals Personal and Ubiquitous Computing and Advances in Human Computer Interaction. His research interests lie in the area of usability and evaluation of mobile systems, in particular context-aware systems, novel text input methods and visualisation on hand-helds.

Gareth Evans

Gareth Evans is currently a Senior Lecturer in the School of Informatics, University of Manchester (formerly the Department of Computation, UMIST). His current research interests include access to computer-based information for people with print impairments; sensory stimulation systems for people with profound and multiple disabilities; and speech synthesis for minority languages. This work has led to the development of a number of commercial systems.

James R. Fruchterman

A technology entrepreneur, Jim Fruchterman has been a rocket scientist, founded two of the foremost optical character recognition companies, and developed a successful line of reading machines for the blind. Fruchterman co-founded Calera Recognition Systems in 1982. Calera developed character recognition that would allow computers to read virtually all printed text. In 1989 Fruchterman founded Arkenstone, a 501(c)(3) non-profit social enterprise, to produce reading machines for the disabled community based on the Calera technology. Following the sale of the Arkenstone product line in 2000, Fruchterman used the resulting capital to expand Benetech, with an explicit goal to use the power of technology to serve humanity. Benetech has numerous social technology ventures, including Bookshare.org, an on-line electronic books library for people with disabilities, and the Martus Human Rights Bulletin System. Fruchterman has also been active in public service, with two stints on U.S. federal advisory committees on disability issues. Fruchterman has received numerous awards for his work, including the 2006 MacArthur Fellowship and the Skoll Award for Social Entrepreneurship in 2004 and 2006. Fruchterman was named as an Outstanding Social Entrepreneur 2003 by the Schwab Foundation and, as such, has become a regular participant in the World Economic Forum in Davos, Switzerland since 2003. Fruchterman believes that technology is the ultimate leveller, allowing disadvantaged people achieve more equality in society.

Han Leong Goh

Han Leong Goh received his Bachelor degree in Engineering in 2002 from the National University of Singapore. He joined the Department of Engineering, National University of Singapore in 2002 as a PhD student and his current research interests include mobile computing and sensor networking.

Marion Hersh

Dr Hersh is currently a Senior Lecturer in the Department of Electronics and Electrical Engineering, University of Glasgow. She previously held an appointment as a Post-doctoral Research Fellow (1982–1985) at University of Sussex.

Her current research interests fall into three main areas:

- Assistive technology for deaf, blind and deafblind people.
- The application of fuzzy and soft computing techniques to sustainable design and decision making.
- Ethics and social responsibility issues in science and engineering.

She has run the successful conference series entitled 'Vision and Hearing Impairment: Rehabilitation Engineering and Support Technologies for Independent Living' with funding from the EC High Level Scientific Conferences Programme jointly with Prof. M. Johnson of the University of Strathclyde. Current research projects in the area of assistive technology include the development of principles of sustainable and accessible design, open and distance learning for people with mobility impairments with colleagues from Germany, Hungary and Romania and intelligent audiometry with her research student Roddy Sutherland.

As well as research, Dr Hersh also has well developed teaching activities in the area of assistive technology, including an undergraduate module written together with Prof. M.A. Johnson. She also regularly supervises undergraduate projects in assistive technology and is looking to market or otherwise distribute an alarm clock for deafblind people developed through these student projects.

Other recent research projects have included the Application of Fuzzy Methods and Other Soft Computing Techniques to Sustainable Design and Decision Making in collaboration with Dr. I. Hamburg, Institut Arbeit und Technik, Wissenschaftszentrum Nordrhein Westfalen, Germany and Mr L. Padeanau of the University of Craiova, Romania. Dr Hersh has recently completed a book on *Computational Mathematics for Sustainable Development* also published by Springer Verlag. Dr Hersh is a Member of the Institute of Mathematics and its Applications, the Institution of Electrical Engineers (UK) and the Institute of Electronic and Electrical Engineers (USA).

Rüdiger Hoffmann

Rüdiger Hoffmann studied Radio Engineering at the Dresden University of Technology and worked from 1971 to 1982 in the electronics industry. He finished his Dr.-Ing. and Habilitation theses in system theory. In 1982, he started his research work in speech communication at the Dresden University of Technology. In 1992, he was promoted to the Chair for Speech Communication. He is now the director of the Laboratory of Acoustics and Speech Communication.

His teaching and research topics are signal and system theory with a special focus on speech processing. He is involved in different research and industrial projects. The main research projects of his group include the participation at the speech translation project Verbmobil (1991–2000), the Dresden Speech Synthesis System (DRESS) and the unified experimental system for speech synthesis and recognition (UASR). The industrial application of the results is mainly directed to the inclusion of speech technology in embedded systems. He has written two textbooks on signal processing (in German) and made numerous contributions to scientific conferences and technical journals. He is the series editor of the "Studientexte zur Sprachkommunikation" (Lecture Notes in Speech Communication).

Brian S. Hoyle

Brian Hoyle is Professor in Image and Vision Systems in the Institute of Integrated Information Systems at the University of Leeds, UK. His interests span a number of areas that critically link sensing systems with integrated processing to yield innovative advances. An example was a monitoring system for a sub-sea robot-welding system for which he was awarded the IEE Measurement Prize. Further major work has been in spearheading the new technology of industrial process tomography (PT); he is a leading member of the world leading research group:

the Virtual Centre for Industrial Process Tomography (www.vcipt.org). He has published his work widely and has gained large scale international and national government and industry funding support. Work has also included the application of ultrasonic sensing in spatial sensing for aids for visually impaired people.

Knowledge/technology transfer has been a continuous thread in his work. He is founder of three spin-out companies: in component mapping systems—*Graticule* (*www.graticule.com*); in process sensing using electrical PT *Industrial Tomography Systems* (*www.itoms.com*); and in ultrasonic sensory aids for the visually impaired, *Sound Foresight*. Sound Foresight was awarded the BBC "Tomorrow's World Award for a Health Innovation" in 2002, for the revolutionary "UltraCane" that provides tactile information of potential hazards to its user (*www.UltraCane.com*). Its engineering was awarded the Sony-sponsored European Electronics Design Application Award in 2003 and the Horner's Award in 2005.

Gunnar Jansson

Gunnar Jansson took his PhD in Psychology in 1969 at Uppsala University, Uppsala, Sweden, on some topics in visual event perception. He has been employed at the Department of Psychology, this university, as research assistant, research associate, and assistant research professor (1957–1979) as well as associate professor (1979–1997). In 1984 he was visiting associate professor at Department of Psychology, Cornell University, Ithaca, NY, and he has been visiting researcher at this university as well as at the Smith-Kettlewell Institute of Visual Sciences, San Francisco, CA, several times. Gunnar has made frequent visits to research laboratories in Europe, USA, Australia and Japan. He has guided some 100 students on different levels in their research tasks, from undergraduate papers up to PhD theses. Since 1997 he has been emeritus but continues with both research and teaching.

Gunnar's research after his PhD thesis has mainly been directed towards basic and applied problems in connection with substitution of vision with other senses for the visually impaired, especially concerning aids for orientation and mobility, tactile pictures and haptic displays. In one EU project where he participated (MOBIC) the aim was to develop an orientation aid for visually impaired people, in another (Pure-Form), to develop a haptic display aiming at presentation of virtual copies of museum statues for haptic exploration. A third EU project (ENACTIVE) tries to find ways of increasing the haptic interaction between the computer user and computer, and a fourth (MICOLE) aims to develop multimodal aids for inclusion of visually impaired children. Typically, Gunnar's research problems concern theoretical analyses and empirical experiments on what information is important for people's solution of the tasks and on evaluations of available technical products and prototypes. His work has been reported in about 75 peer-reviewed journals and books, as well as in a large number of contributions to international conferences and reports in Swedish in different contexts. He has co-edited a book on visual perception theory, G. Jansson, S.S. Bergström and W. Epstein, (1994) (eds) *Perceiving events and objects* published by Erlbaum, Hillsdale, NJ, USA.

Tai Fook Lim Jerry

Tai Fook Lim Jerry received his degree in Bachelor of Electrical Engineering at the National University of Singapore in June 2005. He has since joined the education sector, where he is teaching at a public school. His current interests include J2ME applications and database development.

Michael A. Johnson

Professor Johnson's academic career has concentrated on control engineering, theory and applications. He has significant industrial control applications and research experience. He is the author and co-author of books and technical papers on power generation, wastewater control, power transmission, and control benchmarking. He has had a number of student projects in the area of assistive technology and has successfully trained over 40 engineers to PhD level in the last 20 years. He is joint Editor to the Springer-Verlag London monograph series *Advances in Industrial Control* and to the Springer-Verlag London *Advanced Textbooks in Control and Signal Processing* series. He was an Associate Editor for AUTOMATICA in the years 1985–2002.

His interest in Assistive Technology came through undergraduate teaching commitments which included student group study dissertations at second year level and assorted project supervision at Masters and undergraduate level. These were used to explore some EEE aspects of assistive devices for the hearing impaired. Subsequent collaboration with Dr. Hersh at the University of Glasgow led to a set of targeted activities in this area. These included a new course module, a new conference series and some research projects.

Professor Johnson retired from academic life in 2002 and was made an Emeritus Professor of the University of Strathclyde in 2003. Emeritus Professor Johnson is a member of Deafblind UK and the British Retinitis Pigmentosa Society.

David Keating

Since 1990, David has been head of the National ElectroDiagnostic Imaging Unit at Gartnavel General Hospital in Glasgow and is employed as an NHS Consultant Medical Physicist. He is also an Honorary Reader with the University of Glasgow. Prior to working in vision science, he spent some time working on signal processing systems in cardiology and also in the development of synthetic speech aids for patients with acquired speech loss for which he was awarded a PhD in 1988.

Current research interests lie in the fields of anatomical and functional imaging of the visual system. Recent book chapters and publications have concentrated on multifocal electrophysiology and optical coherence tomography systems.

He is a member of the International Society for Electrophysiology of Vision (ISCEV) and the newly formed British Chapter (BRISECV); the Association for Research in Vision and Ophthalmology (ARVO); the Institute of Physics and the Institute of Physics in Engineering and Medicine.

Charles LaPierre

Charles LaPierre is the Chief Technical Officer and co-founder of Sendero Group LLC. He has a Bachelor and Masters degree in Electrical Engineering from Carleton

University, Canada. His work includes the integration of GPS and dead-reckoning technology. He is one of the inventors of Strider, the first talking GPS for the Blind which evolved from his fourth-year Bachelor Engineering project (1993), and is one of the patent holders for Strider. He has been a software engineer for a number of adaptive technology companies including VisuAide, Arkenstone, Benetech, and Sendero Group LLC.

Roger Lenoir

Roger Lenoir, E. M. MSc, Sound & Music Technology, was born in 1972, Geleen. Roger Lenoir studied Music Science at the University of Amsterdam and Music Technology at the High School of the Arts of Utrecht (Hilversum). From his second year he focused on the analysis and development of abstract music representations systems with special attention for the relation between representation and performance of music. In 1998 he received his BA in Music Software Development (summa cum laude) and Electronic Music Composition. In 1998 he also received his European Media MSc Sound & Music Technology from the University of Portsmouth focussing on Music Software Development (summa cum laude) with his thesis on "Representing musical expression in computer systems". In 2000 he became a member of the research department of FNB Nederland (now DEDICON), where his activities focused on the development and implementation of new computer models of sound and music for provision of (musical) information to vision and print impaired users. Since his appointment to this positions he has been working towards the notion of fundamentally accessible frameworks for information processing that incorporate various layers of abstraction into manageable and practical systems. One of the perspectives Roger has on the field of (fundamental) accessibility is his artistic view. Next to a programmer and system architect, Roger is an artist expressing himself through drawings, computer generated graphics, music and poetry. He is also a founding member of the Openfocus Institute in Amsterdam.

Michael May

Michael May has been a pioneer in new product and business development in a variety of industries for over 25 years. He was on the founding team that invented the world's first Laser Turntable; he invented portable heat devices for medical and sports markets. Mike was the Vice President of Arkenstone where he was instrumental in the development of the Strider GPS project, beginning in 1994. He began two adaptive technology companies, CustomEyes Computer Systems and Sendero Group.

Mike May founded the Sendero Group in January 2000 to make location information accessible to blind people worldwide on laptops, PDAs and mobile phones. Three U.S. presidents have recognized him personally for his efforts in developing and promoting technology and sports for people who are blind or visually impaired. Vice-President Gore said at the White House in 1996 "It won't be long before we see Mike and others wearing GPS devices on their wrists." In the fall of 2001 Mike May and the Sendero Group was awarded a five-year, $2.5 million

Research and Development Grant called "Wayfinding Technologies for People with Visual Impairments: Research and Development of an Integrated Platform."

In addition to product development, Mike May lectures extensively on wayfinding technology and on the scientific and psychological aspects of vision recovery, having been totally blind from the age of three and then regaining low vision at age 46 due to stem cell and cornea transplants. He is the subject of Crashing Through by Robert Kurson, a story of risk, adventure and the man who dared to see.

Stuart Parks

Stuart Parks is currently a principal clinical scientist with North Glasgow NHS Trust. There he provides these specialist services to six Scottish Health Boards. His research interests mainly involve the investigation of spatial and temporal processing mechanisms in the visual pathway. He is currently conducting research into retino-toxic effects of GABA-ergic antiepileptic medication, new methods for imaging the effects of retinal vein thrombosis and the development of cortical and retinal imaging techniques for paediatric ophthalmology.

Stuart is a member of the International Society for Electrophysiology of Vision (ISCEV) and the newly formed British Chapter (BRISECV); the Association for Research in Vision and Ophthalmology (ARVO);nd the Institute of Physics.

Kok Kiong Tan

Kok Kiong Tan received his PhD in 1995 from the National University of Singapore (NUS). Prior to joining the university, he was a research fellow at SIMTech, a national R&D institute spearheading the promotion of R&D in local manufacturing industries, where he has been involved in managing industrial projects. He is currently an associate professor with NUS and his current research interests are in precision motion control and instrumentation, advanced process control, controller autotuning and general industrial automation.

Dean A. Waters

Dean Waters gained a BSc in Ecology from the University of East Anglia in 1988. He worked for the RSPB and National Trust on wildlife conservation until starting a PhD at the University of Bristol on the interactions between echolocating bats and their insect prey. This was followed by a post-doctoral position on echolocation signal processing using artificial neural networks. In 1995, he started a lectureship at the University of Leeds where he conducts research into bat echolocation and bioacoustics. Dean is a director of Sound Foresight Ltd, a University of Leeds spin-out company which develops ultrasonic guidance systems for the visually impaired.

Index