Schriftenreihe

Alistair D.N. Edwards, András Arató,
Wolfgang L. Zagler (eds.)

Computers and Assistive Technology ICCHP '98

Proceedings of the XV. IFIP World Computer Congress, 31 August – 4 September 1998, Vienna/Austria and Budapest/Hungary

Österreichische Computer Gesellschaft
(Austrian Computer Society) on behalf of the
International Federation for Information Processing

Computers and Assistive Technology ICCHP '98

Proceedings

SCHRIFTENREIHE DÉR
ÖSTERREICHISCHEN COMPUTER GESELLSCHAFT
BAND 118

Printed with support of the Austrian Federal Ministry of
Science and Transport.

XV. IFIP World Computer Congress
Vienna, Austria / Budapest, Hungary
31 August 1998 – 4 September 1998

Published by Austrian Computer Society (OCG) on behalf of the
International Federation for Information Processing (IFIP)
Wollzeile 1-3, A-1010 Vienna, Austria
http://www.ocg.or.at

Printed by Riegelnik
A-1080 Vienna, Piaristengasse 19

ISBN 3-85403-118-1 Österreichische Computer Gesellschaft

Computers and Assistive Technology ICCHP '98

Proceedings of the XV. IFIP World Computer Congress
31 August – 4 September 1998, Vienna/Austria and Budapest/Hungary

Edited by

Alistair D.N. Edwards
András Arató
Wolfgang L. Zagler

Österreichische Computer Gesellschaft (OCG) 1998

Program Committee

Chair: A.D.N. Edwards
Vice-chairs: A. Arató
W. L. Zagler

Paul Blenkhorn
Geoffrey Busby
Peter Dal Bianco
Francis Destombes
Pier Luigi Emiliani
Hiroyasu Funakubo
Julio Gonzales-Abascal
Manfred Gross
Kari Marie Helle
Joachim Klaus
Willibald Kremser
Jan-Ingvar Lindström
Klaus Miesenberger
Harry Murphy
Gerlinde Renzelberg
Veith Risak
Arnold Schneider
Constantine Stephanidis
Robert Stevens
Paul Trehin
Felix Tretter
Jiri Vasa
Teréz Vaspöri
Wolfgang Wünschmann

ACKNOWLEGDEMENTS

MAIN SPONSORS OF THE XV. IFIP WORLD COMPUTER CONGRESS 1998

as of 18 July 1998, in alphabetical order

Austrian Federal Ministry of Science and Transport

City of Vienna

European Union

Foundation for Technological Development of Industry

Hungarian Government

IBM Austria, IBM Hungary

ORACLE Austria, ORACLE Hungary

Sun Microsystems

Vienna Tourist Board

SPONSORS

Austrian Federal Ministry for Education and Cultural Affairs; Hungarian Telecom; Hungarian Tourist Board; MATÁV, Hungary; National Committee for Technological Development, Hungary; Post & Telekom Austria; Public Foundation for Open Vocational Training, Hungary; SAP Austria; Siemens Austria; SONY DADC Austria; Taverna, Hungary

COOPERATING ORGANIZATIONS

CEPIS; European Oracle User Group (EOUG); Forum Info2000; Vienna University of Technology

Special thanks to all IFIP and CEPIS member societies that supported the organizers in preparing this event.

OFFICIAL CARRIER

AUA - Austrian Airlines

MALÉV, Hungarian Airlines

Introduction by the IFIP President

I take great pleasure to welcome you to the 15th IFIP World Computer Congress on "The Global Information Society on the Way to the Next Millennium".

This Flagship event of the International Federation for Information Processing (IFIP) is a product of the active involvement of hundreds of IT specialists and many national computer societies around the globe. Ist program and agendas are targeted at topical areas of information processing intended to catch your interest and stimulate debate to suit the needs of IT professionals from academia and industry alike.

The debate will aim to disclose the current state-of-the-art and to project the trends for new developments, which will transcend the threshold of the 21st century and lead society in the magic future of technology.

Moreover, the venues of this IFIP Congress are in the heart of Europe. Vienna and Budapest are historical, cultural and technological crossroads of achievement and civilization. These beautiful cities host the IFIP Congress which in turn provides a forum and a platform for people to meet and discuss issues of common interest and concern.

I hope that you will benefit from this exceptional IFIP Event for your daily work and that you have opportunity during the Congress meeting old friends and making many new ones.

Kurt Bauknecht,
University of Zurich,
President of the International Federation for Information Processing - IFIP

Introduction by the Chairman of the International Program Committee

I am proud to be able to present this technical program. Out of more than 700 contributions from 86 countries, the program committees of the 7 conferences have carefully selected 350 papers, which are presented at the conference and printed in this proceeding

The conferences which you are allowed to switch from one to another cover the main technical aspects of the global information society, like basic theories in the Fundamentals and IT&Knows conference, the hot topics security and legal conditions using net and computers, and important applications such as teleteaching, telecooperation and the use of computers by people with special needs. In the keynote sessions, outstanding experts will introduce the fundamental technologies which made and will continue to make information society possible.

The congress is accompanied by a set of panels, tutorials and workshops. I would like to draw your special attention to the Youth Summit. It will be for the first time that young people will have the opportunity to articulate their views and expectations to the IT community and politicians.

The program gives a good opportunity to study the major technologies and applications which will be the base of future society and to discuss its impact with colleagues from more than 80 countries of our globe.

I hope you enjoy the program, which was so carefully assembled by about 200 ladies and gentlemen in the various program committees, and that your attendance at the congress will have a positive impact on your professional work.

Furthermore, I am sure, that this proceeding will gain a good position in your personal technical library.

Egon Hörbst
Siemens AG Munich,
Technical University of Munich,
International Program Committee Chair

Preface

This fifteenth World Computer Congress is a historic event and it is entirely appropriate that the Sixth International Conference on Computers Helping People with Special Needs (ICCHP) should form a part of that congress. The advances intechnology since the first World Computer Congress are well known. Accompanying the technology development have been changes in attitudes and it is easy to suggest that if one looked at the proceedings of the early congresses, current attitudes to Information Technology were completely unanticipated. The technology is ubiquitous and is often treated with resignation: 'Love it or hate it, it is here to stay.' On the one hand, what is now possible in terms of computing power, was beyond the dreams of most people, on the other the *way* it is used is often the source of frustration and discontent.

Yet there is one area of application of the technology which was both little anticipated but almost universally approved – that of information technology as a means of assisting people who are identified a shaving disabilities or special needs.

The very title of the IFIP organization reflects an early recognition that the technology is concerned with the processing of information, but it must have been hard to realize how fundamental that ability is. It is because information is so important and so flexible that machines that help process it have such a broad range of applications for people with disabilities.

The range of contributions in these proceedings reflect that breadth while we are naturally confined by the current state of the art in the technology. Visual representations dominate mainstream technology, but increasingly it is feasible to implement non-visual alternatives. So it is that contributions on applications for people with visual disabilities have been most common. Among the newly emerging technologies, voice input has (it seems) suddenly become viable and that is also apparent in the submissions received.

A conference such as this can sit in an uneasy position between being adisplay of academic achievement and the channel for communication at a more practical level. I hope we have got the level right. One device that the ProgrammeCommittee introduced was the *Poster with Short Presentation*. This we hoped would give contributors the chance to demonstrate very practical work in an interactive manner, but also to explain the work and theory behind it more formally. Those contributions are also represented as short papers in these proceedings.

A conference such as this – and its accompanying proceedings -do not happen without a tremendous amount of work by many people. The members of the Programme Committee are listed elsewhere in these proceedings and their contribution of time, effort and expertise must be acknowledged gratefully. The programme and the proceedings would not have emerged without long hours of work by the staff at the Austrian Computer Society and the John von Neumann Computer Society and I particularly would like to mention Wolfgang Hawlik and Daniela Poetzl.

I wish to avoid writing anything here that I will regret at a later date when some future ICCHP chair looks back on these proceedings. The safest course then is not to try to make any predictions. Instead let me hope that this work will be looked on as a small but significant step along the way to the future – what ever it may be.

Alistair Edwards
Programme Chair

Contents

Session III

Non-visual Communication for Blind People

Session IV

Access to the World Wide Web

Medicine & Gerontechnology

Session V

Session VI

Session VII

Mathematics

Human-Computer Interaction (1)

Session VIII

Access to Graphical User Interfaces

Human-Computer Interaction (2)

Session IX

Tele-learning & Teleworking

Human-Computer Interaction (3)

Session X

Communication

International

Posters

Universal Support, Opportunities and Challenges

Professor Alan F. Newell, F.R.S.E.

Department of Applied Computing, The University of Dundee, Scotland,
tel: +44 1382 344144, fax: +44 1382 325509, email: afn@computing.dundee.ac.uk

Keywords: *Disability Research, Methodologies, Elderly People, Priorities.*

Abstract

This paper suggests that there is a window of opportunity caused by the growing understanding of the importance of research into computers to help people with special needs. It suggests some of the important research challenges, and comments on the infrastructure needs and research methodologies which could assist in realising the full potential of technology in this field. It also recommends that a greater priority should be given to research into how computers can support elderly disabled people, rather than elderly and/or disabled people.

The growing awareness of how computer can helping people with special needs.

The field of "computers to help people with special needs, and/or people with disabilities has now reached a very exciting watershed. Despite many breakthroughs the field has been "in the wilderness" for a number of years, with the perception that it is a very specialised academic backwater, the last refuge of the poor quality hobby researcher. With few exceptions, it is perceived as a commercial disaster area without the levels of profit likely to attract large companies. The field, however, is now beginning to "come of age" and be recognized as very important, not only in its own right, but also for the effect has on computer developments in general. In 1997, Muller et al [3] made the point that "Computer Human Interface (Research and Development) has maintained a long tradition of concern for access to computer systems by persons with disabilities. However, in practice, much of the field's progress in this area has been slow. More often than not the field's attention to this area has been an afterthought: design and access has been based primarily on able-bodied populations rather than users who are disabled".

Although there is still a long way to go, there is evidence that this situation is changing. Within the industrial sector, major manufacturers such as SUN and Microsoft now have "accessibility" programs. Mircosoft now require vendors to meet certain accessibility design requirements, and have made available a software developer's kit to encourage developers to design their applications for use by people who are blind, deaf, or mobility-impaired. {More details of these programs are also available on the Microsoft web page (1998) (www.microsoft.com/enable/).}

This changing situation is good news for all users of computer systems. The importance of all research and development in Computer Human Interfaces taking into account the full diversity of the potential user population was addressed by Newell in his keynote address to InterCHI '93, where the concept of "Ordinary and Extra-ordinary Human Computer Interaction" was developed [see: 5 & 6] Newell made the point that design which takes into account the needs of people with disabilities can produce better design for everyone. A similar view was expressed more recently by SUN Microsystems when they announced their Enabling Technologies Program, with the comment that it is "....driven by the belief that design to meet the needs of users with disabilities can improve the productivity of ALL users", {see Suns web page (1997)}.

Research Challenges

This growing awareness of the needs of people with disabilities has been seen within the USA, particularly in relationship to access to information via the National Information Infrastructure.

The National Science Foundation of the USA have mounted two workshops with the theme of "Every Citizen Interfaces to the National Information Infrastructure [10 & 11].

These workshops laid out a research strategy for the science and engineering community in the USA. The themes from those workshops included:

- Understanding and representing the functional characteristics of ALL potential users, and the development of creative interfaces capable of understanding user diversity in motivation and interests abilities knowledge and experience
- The development of technology that adapts to the users requirements rather than the user having to adapt to the technology, including self adapting and learning interfaces. and interfaces which predict user behaviour.
- Understand what kinds of information each modality is more suited to expressing, and how do people with disabilities (particularly sensory ones) perceive particular types of information. How can we present information in different modalities without changing its meaning (e.g. subtitles), and what are the most effective ways of translating information across modalities.
- How best do we present information most effectively to people with disabilities in the various modalities (e.g. slow/fast speech etc.), and what does this tell us about improving our presentational techniques for ordinary people. Is it possible to provide universal representations of data, and give appropriate "hooks" for those situations were one modality is not appropriate (e.g. icons and pictures, when the user is visually impaired or their eyes are busy).

In addition, I believe that engineers need to take account of:

- What information do people want, and what they may be prepared to pay for. How do we assist people with disabilities in the navigation of large data bases

- How do we present information through very narrow bandwidths. For example even non-visually impaired users find that accessing data solely using a computer screen essentially gives the user severe tunnel vision, poor resolution and only a limited dynamic range of intensity. How much worse this is for a visually impaired user.

Other more detailed research challenges include the need to develop:

- Interfaces which are not gender biased nor demand a high level of computer literacy, nor good language abilities and which can cope with the variety of age, class, education, language, ethnic origins and abilities which are present in society, particularly bearing in mind demographic and medical trends.

- Remarkably easy to use interfaces which facilitate storage, communication and retrieval of information, browsing and data mining, for a wide range of data.

- Interfaces which automatically adapt, and learn the behaviour of users and what they want to do with their systems.

- Prostheses for remote interpersonal communication. For example, in what ways can email be improved and made more efficient and effective for people with special needs

- Cognitive prostheses - this is an immensely difficult problem, and one for which there is a very great need. A diary is, of course, a cognitive prosthesis, but not a very good one. Elderly people are a very rapidly growing portion of the population for whom memory prostheses are particularly important. As the world gets more complex, and more reliant on vast amounts of data, however, we will all need good cognitive prostheses, not just the elderly.

There is a clear perception of a great need for research in this field in the USA, and there is the TIDE initiative, and other national programs in other parts of the world. There are more than enough research issues in this field, but we should also bear in mind that different cultures have different research methodologies. For example, the European and North American models of research can differ very greatly, (Europe, for example, offers a more holistic approach less dominated by counting and measuring, but the USA is more interested in developing exciting innovative approaches). It is thus important for a fully international approach is taken in this next stage of the development of our field and this international congress will make an important contribution.

Infrastructure Needs

There are certain infrastructure needs to support these activities, and one is dissemination of information. In our field that there are a large number of sources of data, not all of them reliable and it is not easy to obtain a picture of the whole field. We need to bring together this information, but if a data base is too broad it is useless, if it is too narrow you need to access too many different data bases to obtain appropriate information required. In Dundee we have recently been asked to provide for the UK an information service for technology to support students in Higher Education with disabilities. A major activity of the DISinHE (Disability and Information Systems in Higher Education) is the provision of a web site which will provide such information in a structure easy to use form, via a web site which will be fully accessible [1]. We are aware of the dangers of duplication, but an advantage of a Web site is that it is relatively easy to cross-reference other sources of data, and we intend to take full advantage of this. We intend collaborate with other providers, such as the TRACE Center Web Site, to ensure that there is appropriate linkage between data bases in this field and any overlap between data on different sites is necessary and essential.

Research Methodology

It also need to increase the quality of the research done in the field. This can be done in a number of ways. We need to be aware, and take advantage of the leading edge issues in these enabling technologies. Assistive Technology research can often be the first to take advantage of the early manifestations of new technologies. There are many examples such as hearing aids being one of the first commercially successful examples of miniaturization, speech recognition and synthesis systems were, and in many cases still are, used much more widely as Assistive Technology rather than in mainstream application areas, and many AAC devices are a very fruitful application area for Natural Language Processing Research [9].

Researchers must be sensitive to the needs and wants of the current generation of people with disabilities, and this should illuminate their activities. The main responsibility of the research community, however, is to the next generation of people with disabilities. Particularly in this field, there is a tendency to avoid leading edge issues, because these are thought to be of little practical value, or to involve new, expensive and/or untried technology. We do a great disservice to disabled citizens of the future by not giving such work a high priority. People with disabilities have as much right to eventually reap the benefits of Blue Sky research as able bodied people, and a number of the more successful ideas which have come from my research group, and other groups in the world were very much less than popular with the majority of clinicians and users in the early stages of their development than they subsequently became. Some were positively pilloried in these formative stages [4 & 8].

A User Centred Design methodology is absolutely vital in the field of Assistive Technology. The involvement of clinicians and potential and actual users is very important, but this needs to be done with care. At Dundee University, we take into account the needs and wants of users, but this is often greatly modified by our own visions of what can be achieved by Assistive Technology [7] We have a number of panels of users with disabilities who have contributed greatly to our research. Some will have eventually been provided the commercial outcomes of the research to which they contributed, but this can be a number of years after the particular piece of research in which they were involved. We also have two non-speaking researchers [2], who, for various reasons, have a volunteer status. We try to ensure, however, that our research is generic, but informed by individual users, rather than focused on the particular needs and wants of particular individuals. This is a model I can strongly recommend to research groups.

Elderly People

The field also needs to decide priorities for research areas. An analysis of the papers presented to one of the sections of the recent TIDE conference, Newell [10] noted that the spread of papers did not reflect the statistics of prevalence of disabilities: the most striking case being the lack of attention to hearing impaired people, and the amount of work focused on the needs of visually impaired and blind people. To some extent this may be because acoustically amplifying hearing aids tend not to be considered as "Assistive Technology". Also there is a sense in which blind people in "cyberspace" are analogous to wheelchair users within the built environment: a very important group, which are obviously and extremely disadvantaged within the environment, but with a potential danger that they can be mistakenly thought of as the only ones who need to be considered by researchers and developers. A little more disturbing is that, although there are a number of projects investigating "smart houses", there is not a great deal of research investigating the special characteristics of elderly disabled people and research and development of technological support which is appropriate for them. This is a very difficult field which is often perceived as being very unglamorous. Nevertheless it is very important application area demographically, and, if we do not ensure that technology plays an important part in supporting the quality of life of elderly people in the future, society will be significantly the poorer. We need to increase the priority placed on research into how elderly disabled people (as opposed to elderly and/or disabled people) can properly be supported by communication and information technology.

Conclusions

The advances in current technology can be used either to support people with special needs, or to put them at an ever greater disadvantage. The papers in this conference show what can be achieved, and the need for this type of research is being recognised by people, governments, and industry. We have the opportunity to provide the technology which can underpin the political will to provide a better quality of life for people with special needs, and we should grasp it with both hands.

References

[1] DISinHE. The Disability Information Systems in Higher Education web site is bringing together research, development, and provision of Information Technology Support for Students with Disabilities in Higher Education, **http://www.disinhe.ac.uk**

[2] A. McGregor, A. & Alm,N., (1992) "Thoughts of a non-speaking member of an AAC research team", abstract in AAC Augmentative and Alternative Communication, 8(2), 1992, p. 153 (presented at the 5th Biennial International Conference of the International Society for Augmentative and Alternative Communication (ISAAC), Philadelphia, PA, USA, 6-13 August 1992).

[3] Muller, M.J., Wharton, C., McIver, W.J., & Laux, L. (1997) Toward an HCI research and practice agenda based on human needs and social responsiblity, Proceedings of CHI' 97, (Atlanta Georgia, March 1997), ACM Press , 155-161

[4] Newell, A.F., (1992), "Today's dreams - tomorrow's reality", Phonic Ear Distinguished Lecture, Augmentative and Alternative Communication , 8 , June 1992 (Decker Periodicals Inc., Ontario, Canada), pp. 1-8.

[5] Newell, A.F., (1993) Interfaces for the ordinary and beyond,, IEEE Software, 10(5) Sept 1993, 76-78

[6] Newell, A. F. & Cairns, I.Y. (1993) Designing for extra-ordinary users, Ergonomics in Design, Oct 1993, 10-16

[7] Newell, A.F., Arnott, J.L. Cairns, A.Y., Ricketts, I. W. & Gregor, P., (1995) "Intelligent systems for speech and language impaired people: a portfolio of researchin Extra-ordinary Human-Computer Interaction: Interfaces for users with disabilities by A. D. N. Edwards (Ed.), Cambridge Univ. Press, 1995 (ISBN 0-521-434-13-0), pp. 83-102.

[8] Newell, A.F., & Alm,.N., (1994) " Developing AAC technologies: a personal story and philosophy", A. F. Newell and European Journal of Disorders of Communication, 29, 1994 (College of Speech and Language Therapists), pp. 399-411.

[9] Newell, A.F., Langer, S., & Hickey, M., (1998). "The Role ofNatural Language Processing in Augmentative and Alternative Communication" (in press), Natural Language Engineering, Cambridge Univ. Press.

[10] Newell, A.F. (1998), "Assistive Technology Research and Technological Development" presented at "Improving the Quality of Life", 3rd. Tide Congress, Helsinki, June 22nd- 24th 1998

[11] NSF, (1996): "Towards an Every-Citizen Interface to the National Information Infrastructure" workshop hosted by the Computer Science and Telecommunications Board of the National Research Council (of America) in Washington 21-23 August 1996 {"More than Screen Deep.", Biermann, A.W. (Chair) Summer 1997 by the National Academy Press, Washington}

[12] NSF (1997): "Expanded Interactions: broadening the notion of human-centered computing". NSF Workshop, University of Delaware May 21 - 23 rd 1997. {A report of this workshop is also to be published by the National Academy Press}

Session I

Tactile Display

EVALUATION OF INTERACTIVE TACTILE DISPLAY SYSTEM

Yoshihiro KAWAI and Fumiaki TOMITA

Intelligent Systems Division, Electrotechnical Laboratory, Japan

Abstract

We have been developing a support system for the visually disabled that actively recognizes three-dimensional objects or environments. This is a total system that has input, processing, and output functions. The visual information is converted into tactile and auditory information, which can be understood easily by the visually disabled. Experiments on the hardware of an interactive tactile display which we have developed, and the interface of this system were conducted for some blind persons. In this paper, we report on the experiment results and evaluation thereof.

1. Introduction

Much research has been done worldwide on support systems for the visually disabled. As virtual reality technologies have progressed in recent years, tactile devices for 2D and 3D worlds have advanced remarkably. One of them is a tactile display [1-8]. Our group has been developing 3D computer vision, and has applied these technologies to some fields. We have been developing a support system for the visually disabled.

There are two senses, the tactile sense and the auditory sense, for the visually disabled instead of sight. The former is superior for understanding relative position in 2D and 3D space and the shape of objects, the latter is suitable for understanding concepts. It is reported that the level of understanding is improved by the combination of the two senses [9]. We have therefore been developing a computer aided system with a multi-modal interface, using tactile and auditory senses (see Figure 1). It is not only an information conversion system but also is one which the user can communicate with.

As an output device, we developed a 3D tactile display that has a digitizer function. Since the tactile display offers familiarity with image data, we selected it. We added the digitizer function so that users can not only be given tactile and auditory information but can also select the required information from the system to make it easier to understand the 3D world. At the present stage, the system functions of vision input, image processing, translation, tactile display, and voice output shown in Figure 1 have been developed.

Some evaluation experiments on the hardware and interface of this system were conducted, and the results are reported here.

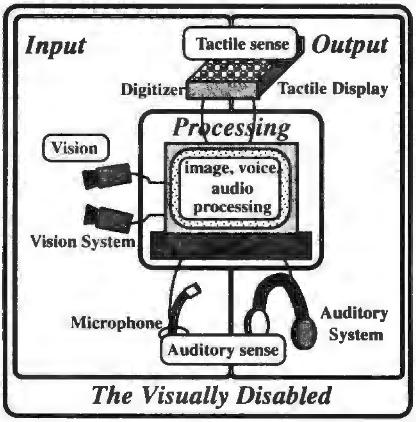


Fig. 1 Conceptual diagram of the support system.

Table 1 Specifications of the tactile display.

Pin arrangement	16 X 16 *pins*
Pin area	175 X 175 *mm*
Diameter of pins	5 *mm*
Spacing between pins	10 *mm*
Height of pins	0-6 *mm*
Shape of pins	[illegible]
Drive	stepping motor
Sensor	tact switch
Size	550(W) X 530(L) x 195(H) *mm*
Weight	28 *kg*

2. Outline of System

An overview of the system is shown in Figure 2 (a) and the structure of the system in Figure 2 (b) [10].

Stereo Camera System

The stereo camera system has two cameras, which are controlled by a personal computer. Each camera is panned and tilted by pulse motors with an angular speed of 50 degrees/sec.

Tactile Display

The tactile display (see Figure 3) represents visual patterns by tactile pins arranged in a two-dimensional lattice. The pin height can be set to several levels to increase the touch information and to represent a 3D surface shape. The major difference is the function of the digitizer. Also, the display has three pushbutton keys for selecting the display mode and hearing voice messages. Table 1 shows the specifications of the tactile display.

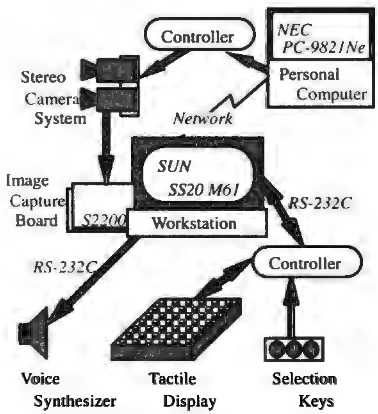
Voice Synthesizer

The voice synthesizer, which is another output device, is used to add more information to the tactile display system, and for reporting the results of recognition and the concepts.

Interactive Interface

Since the resolution of the tactile display is not enough to represent a lot of information at one time, we have developed a multi-level display mode (see Figure 4) consisting of the *Position mode*, *Boundary Shape mode* and *Surface Shape mode*. In the *Position mode*, the user can recognize the relative position of each object by indicating the object position with a single pin. By pushing the voice guide button, the user can hear an auditory explanation about the number of objects. The user can also hear an auditory explanation about the name of the object by pressing the corresponding pin. By pushing the information selection key to change the mode to the *Boundary Shape mode* or *Surface Shape mode*, the user can learn the shape of the object. In these modes, detailed information about the object such as its size and color is explained by voice. In the *Boundary Shape mode*, the user can feel the wire-frame shape, and in the *Surface Shape mode*, scan the convex or concave shape. The user can select one of the three modes easily by pushing the information selection keys, and can recognize each object and the environment by repeating these processes.

(a) Overview



(b) Structure

Fig. 2 Interactive 3D tactile display system.

Fig. 3 Tactile display.

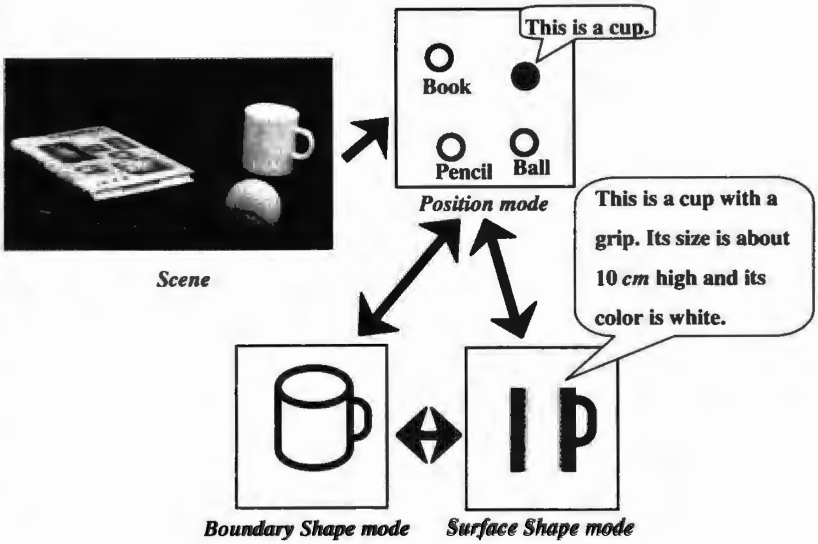


Fig. 4 Concept of interactive interface.

3. Evaluation Experiments

Evaluation experiments of the specifications of the 3D tactile display (*Experiment 1*) and the interactive interface (*Experiment 2, 3*) were performed. The testers were two visually handicapped persons who usually use a tactile map, and three persons who do not. The former two persons were a completely blind man (acquired blindness) and a weak-sighted man whose visual index is about 20 *cm*. The latter three persons were weak-sighted persons who have sufficient eyesight for living.

All subjects were male university students in their early twenties. The purpose and the outline of the system were explained orally first, and after about ten minutes of practice, the experiments were started.

Experiment 1

First, we experimented on the specifications of the hardware of the 3D tactile display. After some practice, the display size, as well as the size, interval, and shape of the pin, etc. in Table 1 were evaluated without voice information. There was no particular problem concerning this. The subjects considered it was a suitable size and shape to scan by finger. They wanted a better display of objects having a detailed shape, namely resolution of length, breadth and height. However, if the object had a simple shape, it was possible to understand enough even at this resolution. It was also felt that the response speed of the pin should be quicker. Generally, many of the comments were limited to the specifications in Table 1.

Experiment 2

Next, various experiments were done concerning the interface in Figure 4. First of all, we tested the subjects' understanding of the shape of a cup and a ball in the shape display modes. For the cup, some patterns were displayed, which were the three shapes viewed from the side, upwards and oblique-upwards in the *Boundary Shape mode* in Figure 5 (a), (b), (c), and one from the side in the *Surface Shape mode* in (d). The users selected the best display, and found that (a) was the best, or (c) was better when combined with (d). The subject who considered that (a) was good wanted a simple expression since it was hard to understand when tactile information became complicated. The subjects who chose (c) said that (a) looked like other objects, however (c) seemed closest to their knowledge of a cup. (d) was not the best, because the display collapsed severely in the direction of height in (d) or the edge parts were hard to distinguish from the smooth changes. We cannot thus decide which display mode is the best.

On the other hand, the *Surface Shape mode* was appropriate for the ball shape. The subjects could not recognize a ball even if voice information was provided, and thought that it was just a circle.

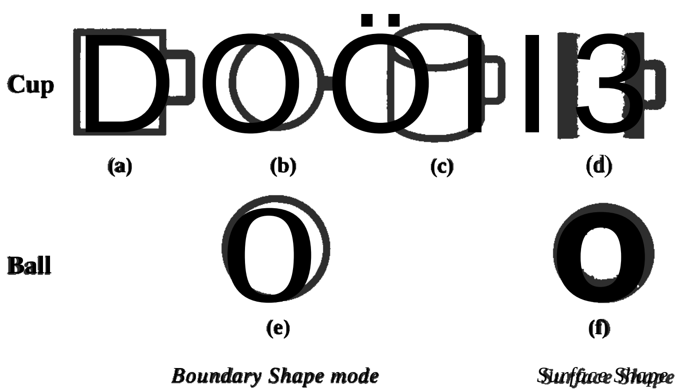


Fig. 5 Two explanations of shape.

(Brightness shows the height of the pin.)

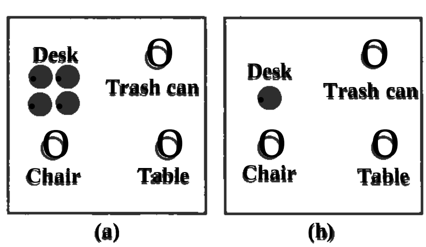


Fig. 6 Two explanations in the *Position mode*.

Experiment 3

We experimented on various aspects of the display method concerning the *Position mode* in Figure 4. As a result, the display from the upper side was chosen, because it was considered most comprehensible similar to a general map. Moreover, we examined whether to project the size of the object. Two kinds of display method, "one pin = one object" and "some pins mean one object", were compared as shown in Figure 6, namely, whether the size of the object should be represented with one or more pins. The result was that the subjects confused the number of objects in the display of the latter, thus the simple method "one pin means one object" was easier to understand.

4. Discussion

The following conclusions were drawn from some simple experiments with the visually disabled persons.

The following points were found concerning the multi display mode.

- The scene can be understood from the three display modes, *Position mode*, *Boundary Shape mode*, and *Surface Shape mode*.
- In the shape mode, the users should be able to select the *Boundary Position mode* and *Surface Shape mode* from their knowledge and preference.
- "One object = one pin" is good in the *Position mode* regardless of the size of the object.

Though the restriction of the hardware of the tactile display has a major influence, as simple a lepresentation as possible is preferable because complex visual information is not understood easily by only the tactile sense. Moreover, issues such as faster response speed, the need for an *Undo button*, etc. were raised.

Concerning the method of displaying the voice information, subjects felt that not only the object name but also more detailed information is necessary in the *Position mode*. However, the output amount was also felt to be useful. It is necessary to note that the amount of voice information should not be too much to remember. Moreover, two subjects, who always use braille, said that it was good to always obtain information by the tactile sense through the braille display because the voice was transitory. It is thus necessary to enlarge the braille display part under the tactile display, but the design must not cause confusion in the tactile sense.

5. Conclusions

We performed evaluation experiments of the interactive tactile display system, and examined the specifications of the display and the method of expressing visual information by the multi-modal method using the tactile and auditory senses. By integrating these senses, it is possible to use voice to make up for the lack of tactile information that cannot be displayed due to restrictions on the hardware. Experiments showed that understanding was improved by the combination. In order to increase the amount of information provided by the combination, it is necessary to conduct a quantitative examination of the ratio of tactile and voice information, and to qualitatively examine the contents of the display and the form, etc. In the future, we will work on improving the response speed of the system, enlarging the braille display part, and developing a display method that can be adjusted to the user's preference, etc.

6. References

[1] Kay L.: Electronic Aids for Blind Persons : an Interdisciplinary Subject, IEE Proc., 131, A, 7, 559-576, 1984.

[2] Koczmarec K. A., Webster J. G. and Tompkins W. J.: Electrotactile and Vibrotactile Displays for Sensory Substitution Systems, IEEE Trans. Biomed. Eng., 38, 1-16, 1991.

[3] Kowalik R. and Postawka I.; The Concept of a Full Screen Tactile Display (FSTD) Driven by Electrochemical Reactions, ICCHP'94, 455-460, 1994.

[4] Fricke J. and Bähring H.: A Graphic Input/Output Table for Blind Computer Users, ICCHP'92, 172-179, 1992.

[5] Fricke J. and Bähring H.: Displaying Laterally Moving Tactile Information, ICCHP'94, 461-468, 1994.

[6] Kawai Y., Ohnishi N. and Sugie N.: A Support System for the Blind to Recognize a Diagram, Systems and Computers in Japan, 21,7, 75-85, 1990.

[7] Shinohara M. et al.: Development of a 3-D Tactile Display for the Blind: System Design, ICCHP'92, 422-430, 1992.

[8] Jansson G.: 3-D Perception from Tactile Computer Displays, ICCHP'92, 233-237, 1992.

[9] Minagawa H., Ohnishi N. and Sugie N.: Tactile-Audio Diagram for Blind Persons, IEEE Trans. on Rehabilitation Engineering, 4, 431-437, 1996.

[10] Kawai Y. and Tomita F.: Interactive Tactile Display System - A Support System for the Visually Disabled to Recognize 3D Objects -, ASSETS'96, 45-50, 1996.

Alternative textured display

BEN CHALLIS & JOHN HANKINSON

-Department of Computer Science

University of York

York, England, YO1 5DD-

TATIANA G. EVREINOVA

-Department of Radio Physics

Rostov State University-

GRIGORI E. EVREINOV

-Lab. for Designing of Information Image Systems

SPECVUZAVTOMATIKA Design Bureau

44/5, 13 Lenin St., 344038 Rostov-on-Don, Russia

E-mail: asni@ns.md.runnet.ru-

Abstract

Texture plays an important role in the display of images and in the forming of subjective impressions for blind people. An electromagnetic textured tactile element has been designed which can emulate static and dynamic states of a surface structure. The surface of a spring has been used as the initial model for a surface unit (texton) where the perceived surface relies on the thickness of the wire used,

the diameter of the spring's coils and the distance between each coil. The controllable parameters are: the density of elements forming a display surface, the frequency of vibration and temperature. These determine the number of possible display states which can vary from 9 (if f = 5Hz, t° = +20°C) up to 57 (if f = 5, 10, 20Hz, t° = +20°C, +30°C, +40°C). A current prototype has an active display area of 15mm X 5mm. The elements of a surface (spring coils) are mechanically connected and do not require any additional interpolation for the display of continuous dynamic objects. Such a TTE design can easily be integrated into a 'line' and therefore into more complex matrices allowing the display of pseudographic and dynamic objects.

Keywords
texture, texton, electromagnetic textured tactile element, pseudographic objects

1 INTRODUCTION

In designing a graphic information display for blind users, it is likely that we (the developers) with normal sight will find it difficult to conceptualise the overall design from the perspective of a blind person. Simply closing our eyes is not a sufficient way of understanding the true depth of this design problem as we heavily rely on our visual based preconceptions of the kind of form an interface should take. Neither is it adequate to expect a blind user to learn to use an interface which is initially more comfortable to us.

If the dialogue between user and computer is based entirely on textual information, then there is no necessity for a graphic tablet. The textual information may be effectively presented on a Braille display or with the help of a speech synthesiser. However, if the user interface also includes graphical objects then navigational problems begin to arise. Therefore, if menu structures, icons and other graphical widgets are to be accessible to a blind user, then a large tablet will be necessary to present these objects and their connections.

As graphic objects can be of a mnemonic, associative or abstract geometrical type, the linear conversion of visual graphical spatial objects to tactile images cannot always be executed through an adequate mode [Kurze, M. & Holmes, E (1996)]. Furthermore, quasi-graphic and complex graphic images heighten the need for a display capable of presenting a variety of types of surface structure. It is important to appreciate that the two best characteristics for the recognition of such surfaces are a high density of structural elements and low frequency vibration [Chavas, M. (1996)].

Presently, the structure of a typical display for presenting graphical information is made up of discrete indication elements which are based on a segmented or matrix structure. The diversity of these displays is governed by the structure of the elements, the model used to display the information, the physical attributes of the tactile screen, including its elements, and the required ergonomic characteristics. It is important to note that in contrast to dot indication elements, matrix elements are capable of displaying more than two states.

Usually, the embodiments of tactile displays are designed primarily to reproduce Braille symbols which are of a textual nature. The dimensions of a whole display are determined, as a rule, from the possible number and sizes of standard Braille symbols, displaying from 40 up to 80 characters per line and up to 25 lines per screen. If an active area is 20in x 15in then the tablet can display the usual format of high resolution screens with a width to height ratio of 4/3 allowing the presentation of one printed page [Fricke, J. (1991)].

The tactile pins of a tablet are set so that they correspond to the points of a rectangular raster. Distances between neighbouring pins vary between 1/10in and 1/30in, depending on both the functional assignment of the tablet and any redundancy allowed by a chosen production technology. The value of 1/10in corresponds to the usual distance between Braille pixels. On the other hand, the resolution of a tablet should be less than 1/10in to allow sensory interpolation of the continuity of a line if all pins of this line are in an identical state.

The largest distance satisfying this condition for pin displays is 1/20in. This value is not ideal as the different movements of pins in horizontal and vertical directions can be felt as uninterrupted lines. However, with a line inclined at 45 degrees, vibration will seem to fragment the line. The line will therefore be perceived as consisting of separate points. This can lead to other problems with line recognition; the user needs to distinguish between a functionally important tactile stimulus and an observed effect due to a low resolution display. Therefore, 1/30in is considered a more preferable distance between pixels for the effective realisation of spatial interpolation forming uninterrupted graphical objects.

A conventional tactile screen based on discrete display elements attempts to form an uninterrupted pattern, in an analogue way, by applying a controlling homogeneous structure or polymeric material to the surface of the display. Unfortimately, these are not widely available for use. An alternative approach to tactile displays relies on displacement converters. Each element contains electromagnetic, piezoelectric, hydraulic or pneumatic actuators and a pin as a tactile object [Kay, L. (1984)]. The number of states for a dot element is usually limited to two, and only rarely encompasses three [Schulz, B. & Rainer, H. W. (1996)]. One exception is a 3D tactile display based on extremely small linear actuators [Shinohara, M., Shimizu, Y. & Nagaoka, H. (1996)].

There is a concept to drive display elements by wires made out of memory metals heated by an electric current. However, the dynamic characteristics of such elements are inferior to electromagnetic actuators. We believe that such an approach, which uses an electrical current to stimulate the user's skin, is inadequate for the synthesis of textured surfaces.

The number of discrete display elements (pins) integrated into a tactile tablet varies over a wide range. Devices such as the 'Optacon' employ the fact that, at any one moment, blind individuals can only feel a small region of a tactile display regardless of its overall size. Therefore, with these devices, the user manipulates an "electronic eye" over the graphical information with one hand whilst a small but variable part of the whole picture is presented to one finger of the other hand. This method reduces the required number of tactile pins to 100 - 150. In contrast to this approach, two German companies have developed a tablet integrating more than 7000 pins with an interval of about 1.2mm. This tablet, used by researchers at the University of Stuttgart, enables simultaneous scanning with both hands. An additional requirement is an input reading device with which the user can scan in a new image [Schweikhardt, W. (1996)]. However, the resolution of this display is too low and the cost is prohibitively expensive.

When graphical information is presented on large tactile tablets, a high cognitive load is placed upon the user due to the large area that has to be explored. As a result, the effectiveness of the user's navigation is reduced and the time taken in forming whole percepts of presented objects increases. The consecutive presentation of dynamically varied objects requires a significant concentration of attention and leads to a loss of presented information. More promising designs, from our viewpoint, are tactile displays that are integrated within multimodal pointing devices [Akamastu, M., MacKenzie, I. S. & Hasbrouq, T. (1995), Yutaka Shimizu, Masami Shinohara and Hideji Nagaoka (1996)], as in these cases the scan surface is simultaneously the display surface, i.e. feedback will be realised by the best of all modes.

To generate a sufficient number of states or levels for presenting complex graphics, dot tactile displays must contain a large number of discrete elements. The purpose of our research was to design an alternative discrete element, the dynamic texton, which would have the following characteristics:

- a controllable surface structure for each indication element.
- variable thermal stimuli and/or frequency of vibration for each element allowing the presentation of a textured surface.
- the number of effective recognisable states would not be less than three.
- the potential to present dynamic and/or quasi-static graphical information by the use of a minimal number of discrete elements.
- to provide a simple method of integration of indication elements into the lines/matrices of a tactile display which can then be used alongside input devices (keyboard, mouse and joystick) for blind computer users.

2 DESIGN

It is possible that the perception of a surface structure occurs in the tactile analyser using the same algorithms through which the visual system decides the analogous task i.e. by processing a temporal structure of the afferent flow and the parameters modulating a density or other characteristic of this afferent flow. In other words, a dynamic analysis of the stimulation of receptors when they interact with the researched object (i.e. a surface) results in the formation of a perceived image of that object, or of its fragments, and to a comparison of this with an appropriate analogue known from a previous experience.

Research into tactile processing showed that the dynamic range for this analyser is narrow in comparison to its visual and auditory counterparts [Sorkin, R.D. (1987)]. This, coupled with the complex interactions between vibro-tactile stimuli, which are in spatial-temporary affinity, has resulted in a fairly conservative approach to tactile display design. However, recent electrophysiological explorations [Anonets, V.A., Zeveke, A.V., Malysheva, G.I. (1992)] have shown that the number of possible 'descriptions' (i.e. states) of an afferent flow at the time of stimulation of tactile receptors can have many more levels than was previously observed (i.e. more than 125).

Tactile receptors, which are functionally specialised by virtue of their anatomy and receptors for touch, pressure and temperature, are distributed around the human body. This, by nature, is an information attribute for the spatial-temporary analysis of the environment. This specialisation divides the sensory fields and extends the dynamic range of perception parameters but does not result in isolation. Quite the reverse, a receptor's specialisation leads to the interaction and integration of sensory flows. Therefore, it seems more probable that the physiological characteristics of the tactile analyser are not low, but the use of inadequate stimuli, including electrical ones, leads to inefficient action on a receptors surface and to a lack of exploration of sensory opportunities.

Motivated by the main problem of the structure discrimination of a surface, we propose to design a tactile display oriented on the separate stimulation of the functionally specific tactile receptors of touch, pressure and, if possible, temperature.

The initial model of the display surface is based on an array of textons. The essential tactile element of a texton is a horizontal spring the surface of which contributes to the overall display surface. The perceived surface of a texton relies on the thickness of the wire used, the diameter of the spring's coils and the distance between each coil. In general the form of a spring is cylindrical but it could equally be rectangular, trapezial etc.. However, the cylindrical form of a spring is technologically superior and satisfies the requirements for the formation of the required surface. The density of surface elements can be controlled rather simply: by stretching or compressing the spring coils and by the speed or frequency of these changes. Both of these are informative attributes in the evaluation of sensations from a presented structure.

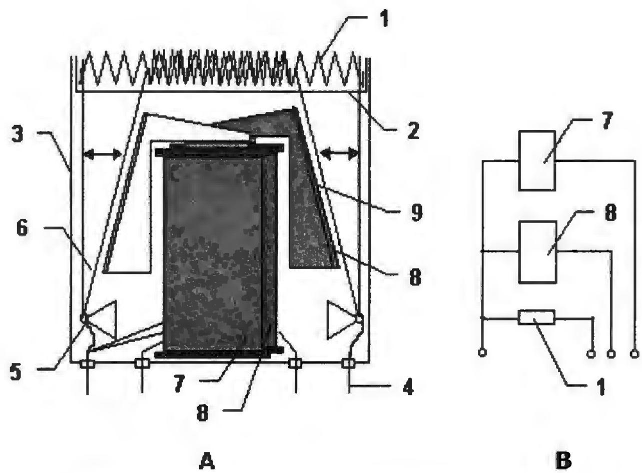


Figure 1 The electromagnetic textured tactile element

1 - a spring; 2, 3 - corpus elements; 4 - electrodes;
5- shafts; 6- levers; 7, 8- electromagnets; 9- layings

The design of such a texture element is limited through technological constraints of achievability, ease of integration and repeatability of the product. These limitations immediately presented restrictions. Electromagnets are used to achieve discrete shifts in a coil's density. Two electromagnets, located as shown in Figure 1, provide four static states of a spring but only three of them are functionally significant when perceiving surface structure (Figure 2 a, b/c, d). These states can be produced via the compression of the spring using one, both or neither of the electromagnets.

A necessary condition for the successful recognition of a given surface is the dynamic displacement of either the receptors or the actual explored surface during the active interaction. In other words the active motion of the receptive surface of the user (fingers) or an appropriate motion of the presented texture must occur. Thus, we can generate some dynamically presented structures whieh are easily recognised. Using electromagnetic switching to manipulate the spring at a frequency of about 5Hz the following dynamic structures can be achieved:

Figure 2 j - simultaneous bilateral periodic compression.

Figure 2 f - alternating compression in opposite phase.

Figure 2 i, h - unilateral periodic compression from one end with the opposite edge remaining free.

Figure 2 e, g - unilateral periodic compression from one end with the opposite edge remaining compressed.

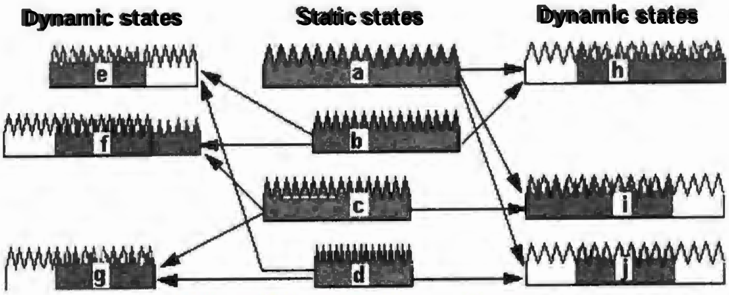


Figure 2 The surface's states of texton

Note that in contrast to the equivalent sensations of the static states b and c (Figure 2) the states i, h and e, g are highly distinguishable. There are three frequency gradations of a texture modulation: 5, 10 and 20Hz which can be effectively recognised at this embodiment. However, more investigation is required into the perception of frequency patterns which depend on the geometrical parameters (thickness of wire, coils diameter and distance between coils) of a texture tactile element (TTE). The design for an electromagnetic texton described above can be integrated into an electrical circuit (Figure 1 B) and would allow the addition of the following functions:

1) a heating control for each TTE.

2) feedback on the contact between the user's sensory surface and the TTE.

The use of high resistance spring materials does not present any technical difficulties. However, the dynamic characteristics required for thermal stimulation are limited. There are at least three recognised thermal levels: +20°C, +30°C and +40°C (±2°C). The realistic use of this parameter would depend on the specific application to which the TTE is to be applied. An improvement in the dynamic characteristics of a TTE can be achieved by additional ventilation of the heated element.

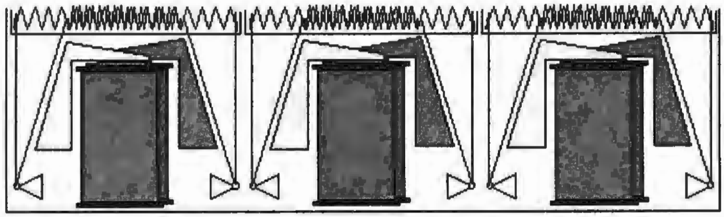


Figure 3

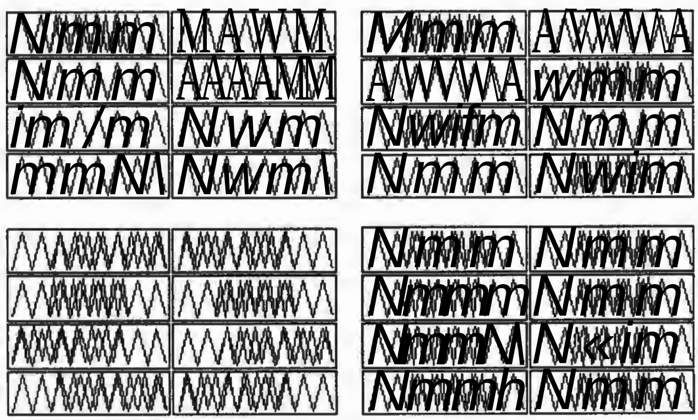


Figure 4

The second function, although not compulsory, could prove useful in the design of the next generation of tactile displays where feedback on user interactions and current placements of receptive fields within a display will be required.

The TTE design presented here can easily be integrated into a 'line' (Figure 3) and therefore into more complex matrices allowing the display of pseudographic and dynamic objects (Figure 4). The active area for a TTE prototype is 15mm x 5mm which permits the easy integration of a single, line or matrix of elements into conventional input devices (mouse, joystick and keyboard). The elements of the surface (spring coils) are mechanically connected and do not require any additional interpolation for the display of continuous dynamic objects.

3 CONCLUDING REMARKS

The electromagnetic textured tactile element has been designed which emulates the static and dynamic states of a surface structure. The controlled parameters are: the density of elements forming a display surface, the frequency of vibration and temperature. The active display area of a prototype is 15mm x 5mm. The number of display states depends on the control parameters and can vary from 9 (if f = 5Hz, t° = +20°C) up to 57 (if f = 5, 10, 20Hz; t° = +20°C, +30°C, +40°C). The elements of a surface (spring coils) are mechanically connected and do not require any additional interpolation for the display of continuous dynamic objects.

The unidirectional polarisation of a texton only permits the display of ordered and periodic structures. The formation of random textures using a similar method (via the controlled compression of a surface) is possible by using new elastic materials and 2D bimorph actuators.

Electromagnetic textons are a promising technology and opportinities exist for further investigation into the development of new projects that incorporate these alternative textured displays.

4 REFERENCES

Akamastu, M., MacKenzie, I. S., & Hasbrouq, T. (1995) A comparison of tactile, auditory, and visual feedback in a pointing task using a mouse-type device. *Ergonomics*, **38**, 816-827.

Antonets, V.A., Zeveke, A.V., Malysheva, G.I. (1992) Possibility of synthesis of an additional sensory channel in a man-machine system. *Sensory Systems*, V.6, No. 4, 100-2.

Chavas, M. (1996) Discrimination tactile de textures, in *Int. Conf. "New technologies in the education of the visually handicaped"* (ed. D. Burger), Colloque INSERM, V. 237, Posters, Paris.

Fricke, J. (1991) A Tactile Graphic Computer Screen and Input Tablet for Blind Persons, in *Proceedings of the First World Congress on Technology for Disabled Persons*, Washington, D.C.

Kay, L. (1984) Electronic aids for blind persons: an interdisciplinary subject, in: *IEE Proceedings*, Vol. 131, Pt. A, No. 1 Sept.

Kurze, M., Holmes, E. (1996) 3D concepts by the sighted, the blind and from the computer, in *Interdisciplinary Aspects on Computers Helping People with Special Needs* (eds.: Klaus, J., Auff, E., Kremser, W., Zagler, W.L.), ICCHP'96, Linz, Verlag R. Oldenburg, Wien München.

Schulz, B., Rainer H.W. (1996) Design and possible applications of a dynamic graphic tactile display, in *Interdisciplinary Aspects on Computers Helping People with Special Needs* (eds.: Klaus, J., Auff, E., Kremser, W., Zagler, W.L.), ICCHP'96, Linz, Verlag R. Oldenburg, Wien München.

Schweikhardt, W. (1996) Interactive exploring of printed documents by blind people, in *Interdisciplinary Aspects on Computers Helping People with Special Needs* (eds.: Klaus, J., Auff, E., Kremser, W., Zagler, W.L.), ICCHP'96, Linz, Verlag R. Oldenburg, Wien München.

Shinohara, M., Shimizu, Y., Nagaoka, H. (1996) Experimental study of 3D tactile display: a step towards the improvement, in *Interdisciplinary Aspects on Computers Helping People with Special Needs* (eds.: Klaus, J., Auff, E., Kremser, W., Zagler, W.L.), ICCHP'96, Linz, Verlag R. Oldenburg, Wien München.

Sorkin, R. D. (1987) Design of auditory and tactile displays, in *Handbook of human factors* (ed. G. Salvendy), New York: Wiley.

5 BIOGRAPHY

Ben Challis was born in Leeds, England in 1965. He graduated from the Leeds College of Music in jazz and contemporary music. Ben completed an MSc in Music Technology at the University of York and is currently a research student in the Dept, of Computer Science at York. His research interest is in non-visual human-computer interaction with particular emphasis on the delivery of musical information to people with visual impairments.

John Hankinson was born in Manchester, England in 1975. He graduated in Computer Science from Cambridge University where he subsequently completed an MPhil in Computer Speech and Language Processing. He is now studying in the University of York Computer Science Dept. His area of research is modelling language structure in music for people with speech and language disabilities.

Tatiana Evreinova was born in Russia in Rostov-on-Don in 1979. She finished secondary school and the music school of the Rostov State Conservatoire. She is now an undergraduate of the Rostov State University in the Dept, of Radio Physics. Her current research interests are alternative and mulitmodal interfaces and HCI.

Evreinov Grigori graduated from the Rostov State University in 1979 where he studied biology and engineering (Computer Lab. of Neurocybernetics (KRINC) Institute). From 1982 up until 1993 he was a leading researcher in the KRINC (Compensatory Behaviour Mechanisms in Extreme Conditions Lab.). He is now a fellow in the Computer Department of RSU and Chief of Lab. for DIIS in

SPECVUZAVTOMATIKA Design Bureau. His current research interests are non-speech audio interfaces, mutlimodal HCI and sensory interaction and substitution.

Hearing Impairment and Speech Recognition

THE VOICE PROJECT (PART 1) GIVING A VOICE TO THE DEAF BY DEVELOPING AWARENESS OF VOICE TO TEXT RECOGNITION CAPABILITIES

Giuliano Pirelli
European Commission Joint Research Centre

giuliano.pirelli@jrc.it
voice@jrc.it
http://voice.jrc.it

Abstract: the difficulties of the deaf go beyond the loss of hearing itself and underline a more general problem of lack of communication. The paper presents an overview of the VOICE Project, a European Commission's Telematics Programme Accompanying Measure. The Project is chaired by the Institute for Systems, Informatics and Safety of the Joint Research Centre, in collaboration with Kepler University of Linz, Software Solutions and FBL software houses near Milan, ALFA and CECOEV Associations of the deaf of Milan, the Institut for the deaf of Linz. The project proposes the promotion of automatic recognition of speech in conversation, conferences, television broadcasts and telephone calls, with their translation into PC screen messages. It also proposes to stimulate and increase the use of new, widely diffused technologies, namely the Internet, with the objective of uniting, by means of an Internet VOICE Forum, Associations, companies, universities, schools, public administrations and anyone else, who may be interested in voice recognition and could benefit from such research.

Keywords: Voice, Speech, Recognition, Communication, Deaf, Deafness, Hearing, Subtitling

1. The Joint Research Centre of the European Commission

The Joint Research Centre is the European Commission's own research centre. It was created to share, on European level, the large investments needed to carry out research on nuclear energy. Over time its tasks have developed into other areas in which a common approach on European level is necessary. JRC provides neutral and independent advice in support of the formulation and implementation of the European Union's policies. In addition, it offers unique training services to individuals and companies and organises workshops for scientific and technical workers in advanced sectors of science.

The activity areas of the Institute for Systems, Informatics and Safety (ISIS) and of its Unit for Software Technologies and Automation (STA) include the innovative application of information and communication technology, dependable software, animation in medical imaging, network multimedia techniques in training and education.

1.1. JRC-ISIS's Exploratory Research Programme

JRC-ISIS's role in 1996 in the previous themes was oriented towards the provision of scientific and technical support to the EU services and initiatives. Moreover, a levy of 6% of the institutional budget was used to finance *exploratory research*. In 1996 the scientific staff of ISIS made a total of 65 proposals. The ISIS Scientific Committee judged the proposals on originality, appropriateness, soundness and cost and produced a shortlist of 16 proposals, 12 of which were then funded. In particular, two projects are carried out by the STA Unit concerning the interface between *Life Science* and information technology to provide help for the disabled and the elderly:
Information technology aids for people with special needs - Voice to text conversion for the deaf
Brain-actuated control - using EEG pattern recognition to help the disabled.

I had the honour of relating to the ICCHP-96 Conference on the starting of these Projects. Since then, they achieved encouraging results and are providing a better definition of the requirements of people with special needs and a more collaborative work between technicians and non technicians, in these interdisciplinary activities.

1.2. The VOICE Project's first steps

JRC-ISIS has undertaken, as from the beginning of 1996, a number of the tasks described later on in more details and related to integrating voice to text recognition into local conversation and telephone conversation for the hearing impaired. The objective was the development of a demonstrator necessary in generating awareness and stimulating discussion regarding the possible applications of voice to text recognition. Technical objective of this research was the set up of a cluster of laboratory prototype applications related to voice to text recognition, intended for any user and particularly for the deaf. This *VOICE Laboratory* included the necessary software, hardware and network capabilities.

Contacts with producers of voice to text recognition systems, research centres, telecommunications firms and television broadcasters, created a coherent overview of the state of art in voice to text recognition, voice analysis and text to speech systems. Regular contacts with the Associations of the hearing impaired gave the opportunity of analysing the special needs, resulting from difficulties in hearing and in speech. Applications of information technology have been considered, in relation to a general problem of lack of communication, in many aspects of the life of the deaf (and in a different way for the blind) and of the elderly.

With the aim of facilitating the contacts and establishing a common goal, JRC-ISIS gave some Associations the opportunity of creating a *VOICE Forum* on the Internet, by allocating space for them on a Web server and providing technical assistance. Since then, the Associations have shown great interest in participating to the Project and they feel reassured by JRC-ISIS' mission, as an impartial European R&D centre, with expertise in innovative applications of information technology. The *VOICE Forum* begins to be a known Internet site and several Associations are adding information to it or communicate their interest in testing the demonstrator and participating to the foreseen meetings and workshops.

Additional software is being created and integrated into the system to turn commercial speech recognition packages into user-friendly programs modelled on the requirements of the users. The technical part of the Project is developed in collaboration with FBL software house, which is experienced in applications of voice to text recognition systems to the disabled. The final operational capabilities of the demonstrator is to achieve a necessary standard of functionality (in order to prove the validity of such applications to companies and manufacturers) in subtitling school lessons, conferences, television transmissions, telephone calls.

2. Applications of voice to text recognition for the deaf

Although voice to text recognition packages are marketed primarily as a means of allowing people in businesses to create documents without using the keyboard, it is an application that holds great advantages for the hearing impaired, blind and physically handicapped, as well as people without special needs.

2.1. State of the art

A great deal of money and man-hours have been invested in developing voice to text products in the last ten years, but the progress only in the last three years has been very noticeable. This is in part due to the wider diffusion of PC's with greater processing power. Voice recognition systems are reaching a very good level of development and begin to be widely available for PCs. They are used by lawyers for preparing drafts that will be read and checked for errors and by radiologists, who do not have their hands free and make use of a very specific dictionary. The software that until now could only recognise words separated by short pauses *(disjointed speaking)*, is being replaced by new releases, which present very significant improvements and recognise continuous dictated text *(continuous speaking)*.

Our interests are concentrated on systems that run on PC's since they are more affordable and appropriate to the final user. In this sector, IBM and Dragon Systems offer systems working in several European languages. Finding solutions and ways of adapting such software for the use of a disabled person is in fact encouraged by this increase in market, affordability and user-friendliness.

A widely used application is the subtitling of television transmissions, very powerful help for deaf people, particularly for the language learning and training for deaf children. The importance of the educational aspect lies in the fact that subtitles are for a deaf child one of the most powerful learning tools of any language, just as a hearing child would learn from things it heard. Similarly it gives hearing impaired adults the opportunity to enrich their vocabulary. Since subtitling of television transmissions is the result of a manual preparation of files to be transmitted in Teletext format, most of the subtitled transmissions are films. Subtitling of live programs and of the news is rarely performed.

Subtitling of conferences, even those addressed to the deaf, is usually not available. Sign language interpreters provide a significant help for the deaf who knows sign language, but other participants or partially hearing impaired, elderly and foreigners are unable to understand sign language. Moreover this activity is *lost* after the conference, being of no use for producing proceedings or abstracts.

In telephone communication, *Text-telephones* have already proved themselves vital from a deaf person's point of view. These systems do, however, present one major problem, that is, all people wishing to contact a deaf person on such a machine must possess one themselves. This makes such a means of communication awkward and expensive, both for the deaf *and* for those they wish to call.

2.2. User needs

The difficulties of the deaf are beyond the loss of hearing itself, and underline a more general problem of lack of communication. Help in reducing the gap between the deaf and the hearing world should be enforced. Automatic recognition of speech in conversation, conferences and telephone calls, with their translation into PC screen messages, could be a powerful help. Please refer to an other paper *(The VOICE Project - Part 3 - The communication needs of the deaf)*, presented by Alessandro Mezzanotte, President of CECOEV, to the *VOICE Workshop*.

Hearing impairment is a particularly important disability to be taken into consideration since it affects people of all ages and is something that often becomes worse with age. It is also important since one of the main forms of modem communication, the telephone, is as yet of no or of very little use to this community for oral communication (while it is useful for the transmission of faxes). Other modem means of communication, although not completely useless, generate frustration by providing only part of the information in a form accessible to them. An example of this is the television which, when not subtitled, supplies very limited information.

In some European countries, it is usual to think that hearing impaired people would had difficulties trying to learn to lip-read and speak and should therefore make use of sign language and attend special schools. In others there is another approach to the problem. In Italy the law encourages the *integration* of deaf children in the normal schools, with a *remedial teacher*, without the use of sign language. Some Associations, like ALFA in Milan, are getting very good results from helping the children following this approach, and do so with children joining primary school right through to those finishing the University and finding job afterwards. Despite the fact that good results are achievable, they demand an enormous effort, which could be greatly reduced through the use of new technologies.

2.3. Market situation and prospects

It is worth remembering that the market of the hearing impaired consists of between 1% and 5% of the population (according to the degree of the hearing loss), which represents millions of people in Europe. This field can be enlarged to take into account also those loosing their hearing, having hearing problems, who can hear but are vocally impaired and even normal hearing people who cannot hear due to the noise in their environment. Moreover, a lack of commiunication similar to that experienced by the deaf also affects the disadvantaged, the people living in foreign environments and the elderly. When united this group consists of more than 30% of the total population.

The new products seem well suited for the needs of the deaf. The modification necessary for some test are of limited extent and could have been foreseen and developed by the producers of voice to text recognition systems, if only they could have the time and the willingness of concentrating on this aspect. But the rapid grow in the voice recognition systems has as a consequence the fact that the experts in this field are very few and they work on the development of other aspects of more immediate use.

Nevertheless this could be an opportunity of a great interest for the producers of speech recognition systems, since the deaf could accept the present limited accuracy of recognition, as a complement to his lip-reading skill. Even the more limited accuracy of recognition over the telephone line, is an interesting starting point for the deaf. The Associations of the deaf are considered both as the most interested and critical user group for all the possible applications in this area, and thus the most motivated for testing a system which will be improved for all users, also in related fields, such as video-telephones or on-line television subtitling in several languages.

The proposed alterations or additions to existing software could be easily added to future releases by the software producers interested in enlarging their targeted market. This will improve the quality of life of persons who at present have difficult access to information and communication.

The proposed demonstrator will enhance a better use of standard products and the definition of new services. The market is ready to accept and spread them, as soon as their quality will be improved and considered good by the users.

3. European Dimension

Hardware, software and services producers of voice to text systems hesitate to invest more, since the user needs are not translated into technical specifications and are sometimes not even known. On the other hand, the Associations of the disabled have a limited overview of possible technical new solutions and rarely have the opportunity to participate in the feasibility studies of new projects. Those who have to take decisions in associations, institutions, political bodies, information technology factories, telecommunications services need for a valid reference point. All the concerned parties look for *Positive Actions*, which might be of specific use to them and an important reference for others.

What lacks is essentially a better definition, from a technical point of view, of the needs of the disabled to enhance collaborative work between technicians and non-technicians. The *VOICE Forum* could play an important role in this field and the European dimension of such a broader co-operation is of great importance, allowing a scale factor for the study and the development of technical aids and ensuring a large impact of the results. This will improve the mobility and the accessibility to information, offering an additional means to participate fully in the information society and improving the quality of life.

There are technical solutions, at a pre-competitive stage, to help the deaf and an effort is required to promote them at EU level, so as to benefit of a large scale factor. Also the care of multilingual aspects should be considered at a European level, since most of the concerned Associations are only at a national level; JRC-ISIS will provide know-how independent of the language. The expertise of the Partners, the previous analysis of the user needs, the availability of laboratories (hardware, software) as well as of demonstrations, the experience in organising meetings and workshops, will help in expanding the present *VOICE Forum* at EU level and to use it as an Internet server for the deaf.

3.1. The VOICE Project - a Telematics Applications' Accompanying Measure

We felt that all the activities started at JRC-ISIS with the collaboration of its Italian Partners, could get a particularly important push if the tests and the dissemination of the results could be organised in several countries. So we enlarged our group, proposing at first to the Institute for Computer Science of the Johannes Kepler University of Linz and to the Institut für Hör- und Sehbildung (IHSB) of Linz to join us.

We created a Consortium of partners with whom we could collaborate on the Project. In order to bring the activities to an broader European level, we prepared a proposal for an *Accompanying Measure*, which we submitted to the Telematics Applications Programme Call in April 1997. The proposal; *VOICE - Giving a VOICE to the deaf, by developing awareness of VOICE to text recognition capabilities*, has been selected and we are at present (April 1998) in the last negotiation phases for starting the Project.

The Consortium proposes to continue and enlarge the activities in this field, and to develop awareness of the capabilities of voice to text recognition systems. The Consortium will play a technical and social role in collecting information and presenting it in a coherent way to the producers of voice to text recognition systems and researchers. The aim is that of disseminating information on how the producers may help the users with disabilities by limited improvement of their standard products and on how the users with special needs may collect useful information and translate it into technical specifications.

JRC-ISIS is acting as scientific and technical co-ordinator of the Project and is developing several specific aspects of the research. FBL software house, which is experienced in applications of speech recognition to the disabled, is developing additional software and integrating it into the demonstrator to turn commercial speech recognition packages into user-friendly programs modelled on the requirements of the users. Each step of the activity is discussed and checked with ALFA and CECOEV Associations of the deaf in Milan. Kepler University examines the Italian results, verifying their validity in Austria and helping IHSB in the Austrian validation phase.

3.2. Objectives and strategic approach

Main objectives of the VOICE Project are: to investigate into voice to text recognition for automatic subtitling of conferences, school lessons, television transmissions and telephone conversations; to spread the use of general purpose voice to text recognition systems and to improve the prototypes developed until now; to demonstrate the prototypes to relevant organisations and in international conferences; to use a *VOICE Forum* on the Internet as a Project tool for collecting and spreading information on technical aids for the deaf.

The VOICE Project proposes not only the promotion of new technologies in the field of voice to text recognition, but also to stimulate and increase the use of new, widely diffused technologies, namely the Internet. The objective of the project is that of uniting, by means of an Internet *VOICE Forum*, Associations, companies, universities, schools, public administrations and anyone else, who is interested in voice recognition and could benefit from such research. The *Forum* will become an intermediary between the different concerned groups and will help in collecting information on the user needs and on the validation of the prototype demonstrator. It will enhance collaborative work between technicians and non-technicians and will help in disseminating the results.

At present JRC hosts and maintains the sites of AFA, ALFA, CECOEV and ENS Associations of the deaf, with information including: Statutes, contact numbers and addresses, meetings, electronic copies of a selection of their newspaper, a research carried out into the hours and accuracy of the television broadcasters, a list of their archive of subtitled videocassettes. The current site provides a very strong foundation on which the creating awareness side of the *VOICE Project* can be built. This is an important part of the Project itself, since it demonstrates, to all those involved, the effectiveness of this means of communication for the deaf community.

All the phases of the Project will be developed with continuous and tight participation of the users. Several European conferences and workshops will be organised in view of helping them to discuss their needs with the industry and services providers: ICCHP-98, Vienna and Budapest, August 98; HANDImatica-98, Bologna, November 98; Linz, first semester 99; JRC-Ispra, second semester 99. The demonstrator will be presented and used for generating prototype live subtitling for the deaf participating to the conferences. The meetings will not only concern the technical aspects, but will also try to bring the manufacturers and producers closer to the users' needs.

The Partners of the Consortium represent different sectors of experience and of activity (research, universities, private IT companies, Associations of the deaf and of their families, Institute for the deaf) and may ensure the complementary skills in order to cover all the aspects of the VOICE Project. ALFA, CECOEV, IHSB, whose members are more than one thousand, represent three different ways of approaching the problems of deafness, due to different culture and language aspects. JRC-ISIS, as an impartial European R&D centre, is in an ideal position to facilitate the dissemination of information and understanding of user-requirements.

The linguistic aspect of the software packages has been considered choosing software packages already available in several European languages. Since most of the new IT packages are produced in English language, JRC-ISIS is testing them in English and the users are doing so in Italian and German, as to cover different linguistic approaches. The acquired know-how will be made available for applications in the other languages. Some contacts have been already established with the University of York and the NDCS Association in UK, French ANPEDA and Belgian APEDAF and TELECONTAC, which showed interest in following the Project. As complement to the VOICE Forum, a *VOICE Special Interest User Group* is being created and will hold its first meeting during the ICCHP-98 Conference in Vienna. It will provide the Project with a larger audience and will participate to the peer review of the deliverables for which this is appropriate.

3.3. Technical aspects of the demonstrator

One of the objectives is the extension of a cluster of demonstrator applications related to voice to text recognition, some of which have been developed on the basis of a multimedia laboratory prototype. The system could be of use for subtitling conferenees, television transmissions and telephone conversations. It involves integrating standard speech recognition software into flexible applications that will help in ensuring low costs and easy use. The technical aspects are described in a second paper (*The VOICE Project - Part* 2) presented to ICCHP-98. In view of an other objective of the Project, which is the *VOICE Forum*, the laboratory will also provide a means of generating and managing Web pages on the Internet, as well as e-mail capabilities.

On the basis of the first experiences, a new prototype demonstrator of automatie subtitling of conferences, based on speech recognition, has been developed. It has been presented in the first quarter of 1998 to some schools that had declared their interest to participate to the Project. The presentation of the Project has been followed by a simulation of a school lesson, with topics on literature, history, world explorations, spatial geography, eleetronics and art, by using the prototype demonstrator for subtitling the speaker's voice.

The prototype will be tested in real situations of use for subtitling school lessons for the benefit of the deaf students. It will visualise the dialogue pronounced during the foreign language lessons, for the benefit of the hearing students, or the lessons of the host country's language for the benefit of any user, particularly the immigrated. Some tests have been also foreseen for subtitling university lessons and printing summaries. The use of the *VOICE Forum* and of the Internet will be encouraged, since this aspect is particularly important for the deaf in order to communicate with his hearing friends for home works and social contacts.

4. Final goal, autonomy and quality of life

The impact that such a Project may have is enormous, changing several aspects of every day life for an important portion of the population. At present, the difficulties in communication maintain the deaf community rather isolated from the world of the others. This demands relevant costs for sign language interpreters or not automatic subtitling. Moreover these services are not available in meetings which are considered less interesting for the deaf, thus increasing the communication difficulties of the hearing impaired and their feeling of being obliged to a few specific fields of interest.

- A wider diffusion of subtitles will greatly increase the interaction of the deaf with each other as well as with the society in which they live. When more conferences, meetings and discussions slowly become subtitled, there will be an increasing in participation from the hearing impaired community.

Once started this improvement of their integration and interaction in the society will have a snow ball effect and it is therefore: this initial push that is so vital.

By an increase in subtitling capabilities, television will become a more useful source of information. The use of subtitles in the telephone calls, which involve everyday communication in society, will greatly increase the interaction of the deaf community. This contribution will increase the effect that their decisions have on the surrounding environment, which will subsequently improve their standard of living. If it will be possible to close a huge gap in the distancing caused by inappropriate means of communication, also the national spending for benefits for the deaf will be reduced. More integration in society and more autonomy in they every day life are the basis for any further improvement. An easier access to schools and universities will allow a more satisfying life and also a better choice of a work corresponding to personal capabilities and, at large, more economic productivity for the society.

The demonstrator will be tested not only on a technical point of view, but also as opportunity for discussing other problems related to the technical ones. The different implications will be discussed with the users, the producers of voice to text systems, television broadcasters, telecommunications firms, etc. in order to see, foresee and understand the problems that will come out in the exploitation of the systems. The Consortium will be in some way at the disposal of the Associations of the deaf, that may contact the developers, as representatives of the needs of a large group of users, and clarify some precise technical points. Thanks to the gained experience, the deaf users should be in a position as to influence, by valid technical results, some aspects of the commercial development of *Voice* products and to convince the services producers of the opportunity of using the newly available products.

We feel that the proposed way of managing the pauses in speech (as it is explained in the aforementioned second paper) gives a very deep feeling of communication between the speaker and the audience. The speaker may so decide at each moment the rate of speaking in function of the audience, of their familiarity with the dictionary, of their being fluent in reading, etc. This proposal is quite different from many other projects, since we do not propose to develop specific software. We just feel that the commercial products will reach good results in the near future and we try to convince the producers to take into account the needs of the deaf. At the same time we try also to help the deaf to get ready to use the systems and explain their expectations to the services providers.

The technical goal of the Project is to develop a prototype with just the basic functions for holding conferences. The final aim is *not* that of developing a final commercial tool, but on the contrary that of using a prototype demonstrator of limited life time (possibly less than the two years' life of the Project) for disseminating awareness so that the producers will include some of the basic functions of the demonstrator into their standard commercial products.

THE VOICE PROJECT (PART 2) HARDWARE CONFIGURATION AND SOFTWARE DEVELOPMENTS

Angelo Paglino
FBL

Giuliano Pirelli
European Commission Joint Research Centre

fbl@lomellina.it
voice@jrc.it
http://voice.jrc.it

Abstract: automatic recognition of speech in conversation, conferences, television broadcasts and telephone calls, with their translation into PC screen messages, could be a powerful help for the deaf. The paper presents the technical aspects of the VOICE Project, a European Commission's Telematics Programme Accompanying Measure. The Project is chaired by the Institute for Systems, Informatics and Safety of the Joint Research Centre, in collaboration with Kepler University of Linz, Software Solutions and FBL software houses near Milan, ALFA and CECOEV Associations of the deaf of Milan and the Institute for the deaf of Linz. The hardware interfaces and the software developed by FBL software house, in collaboration with JRC-ISIS laboratory, are presented.

Keywords: Voice, Speech, Recognition, Communication, Deaf, Deafness, Hearing, Subtitling

1. The VOICE Project technical aspects

Automatic recognition of speech in conversation, conferences and telephone calls, in order to translate the voice into PC screen messages, could be a very powerful help for the deaf. One of the objectives of the *VOICE Project* is the development of a demonstrator necessary in generating awareness and stimulating discussion regarding the possible applications of voice to text recognition. The Project proposes not only the promotion of new technologies in the field of voice to text recognition, but also to stimulate and increase the use of new, widely diffused technologies (such as the Internet) with a particular emphasis on the problems that may be encountered by the deaf.

The *VOICE Project* is chaired by the Institute for Systems, Informatics and Safety of the Joint Research Centre, in collaboration with the Institute for Computer Science of the Johannes Kepler University of Linz, SoftSol and FBL software houses near Milan, Associazione Lombarda Famiglie Audiolesi (ALFA) and Centro Comunicare è Vivere (CECOEV) both of Milan, the Institut für Hör- und Sehbildung of Linz (IHSB).

For a general overview of the Project, please refer to a previous paper (*The VOICE Project - Part 1*) presented to ICCHP-98 Conference by Giuliano Pirelli, the co-ordinator of the Project. For the user needs analysis and validation of the demonstrator, please refer to an other paper (*The VOICE Project - Part 3 - The communication needs of the deaf*), presented by Alessandro Mezzanotte, President of CECOEV, to the *VOICE* Workshop and the *VOICE Special Interest User Group* Meeting, which will be held during ICCHP-98. The present paper presents the technical aspects (hardware configuration and software developments) of the *VOICE Project*.

1.1. SoftSol and FBL Software houses

Software Solutions (SoftSol) is a team of professionals with great experience in networking technologies. The Company has also co-ordinated the development, with Master Soft and FBL, of an application for the blind, to enable the interrogation of various telephone directories with vocal orders. The program makes possible the PBX management by blind person using the Dragon Dictate engine for the vocal input, integrated by FBL, whereas Master Soft developed the speech synthesis for the output. SoftSol is in charge of the financial co-ordination of the *VOICE Project*.

FBL is a qualified distributor of the IBM VoiceType software package and has a long experience as system integrator of Dragon Dictate. With the collaboration of Aries it has gained significant experience in the domain of voice controlled applications for personal computers. In 1992 Aries introduced the first version of Dragon System's dictation software in Italy. This was followed by FBL installing the program on machines for motor deficient users. The company has since then been responsible for hundreds of installations of this type, using the more recent version of Dragon Dictate for Windows. FBL has collaborated with ALFA and ASPHI in giving presentations on the potential of voice to text recognition at several conferences and meetings. FBL collaborates with JRC-ISIS *Voice Laboratory* for the technical developments of the *VOICE Project*.

1.2 Design for all

The *VOICE Project*, according to JRC-ISIS background, TIDE policy and FBL methodology, is developing prototype applications of information technology for people with special needs, using as far as possible hardware and software commonly available on the market. This allows reducing development and maintenance costs, improving the quality of products for the normal market for *any* user, and eliminating new barriers, which often are created by new information technology tools. Also in the case of the *VOICE Project*, the *design for all* products that JRC-ISIS has developed or of which has asked the development to FBL, should have two mains characteristics: ease of use and low price. Experiences show that this goal is possible also with the speech recognition technology, even if at the very beginning this could have seemed too ambitious or just an impossible dream.

The current voice to text recognition packages, produced by companies such as IBM and Dragon Systems, are of very high quality. Their achievements are the result of a huge investment of time, manpower and money. Our goals are to integrate their products into systems for the deaf and to *create awareness* amongst the manufacturers, the hearing impaired community and service providers. This is being done through the creation of demonstrators and prototype systems, introducing the hearing impaired to the available technologies, showing them the possible applications (so that technical specifications can be laid out) and allowing *them* to approach companies and government entities themselves.

2. Market Background

The first experience made in Italy using voice to text recognition systems running on standard PC is dated 1992: Dragon Dictate package (Dragon Systems inc., Newton, MA, USA) was running under DOS on a PC with a i80486 processor and an audio card developed by Dragon. The speed was about 20 to 30 words per minute and the price still too high. At the same time IBM developed a speech recognition system on Risk platform. In 1994 two new packages were announced using the Windows operating system platform: Dragon Dictate for Windows and IBM Voice Type. They had the same basic characteristics: *disjointed speech*, 60 to 70 words per minute, processor i80486, 16 Mb Ram, dedicated or large market audio card, price lower than 1,000 ECU.

At the beginning of 1997, at the time of the preparation of the proposal of the *VOICE Project* for the Telematics Applications Call, there were just a few products for general use, with some limits in their functionality. The most common products that operated on 486 or Pentium PC's and also offered packages in various languages, were those produced by IBM and Dragon Systems. They were available in various languages, which included English, French, Spanish, German and Italian. No other packages had been released with the same characteristics. The linguistic aspect of the software packages is one that had to be considered due to the European dimension of the project.

Dragon Dictate allowed a user to dictate up to 60 words per minute for people who had *trained* the system for a few hours. It also let the user guide the mouse pointer across the screen, by means of vocal commands. This system could alter the last recognised word if it felt that it had not been recognised correctly. IBM VoiceType Dictation allowed a user to dictate up to 60 word per minute and worked in groups of three words at a time, which were displayed as highlighted text prior to their confirmation. A dictionary and a series of probabilities (that one word should be followed by an other) were used in checking that recognised words had been correctly understood. One of the difficulties encountered in the use of these systems related to the yet insufficient quality of the recognition. This was in part due to the necessity for the speaker to insert short pauses in between two words *(disjointed speech)* and also to having to tell the system of punctuation marks.

In July 1997 IBM released in Italy the package Med-Speak for *continuous speech* recognition. At that time it only contained a dictionary for radiologists. Dragon Systems was also releasing the package Naturally Speaking for *continuous speech* recognition. The main characteristics of both systems are *continuous speech*, dictation speed greater than 100 words per minute, high precision (greater than 95%), large dictionaries. The speech recognition engines of both these packages no longer depend on there being pauses between words. This will have a great impact on the quality of voice recognition during conferences and more particularly across telephone lines, since the problems caused by the background noise of the signal during pauses will no longer be present.

3. The VOICE Laboratory

As from the beginning of 1996, a multimedia *VOICE Laboratory* prototype was installed at JRC-ISIS and FBL started working on the Exploratory Research *VOICE Project* of JRC-ISIS. The two main applications on which the research was based were the integration of voice to text recognition software into a subtitling system for meetings, conferences, lectures and television, and a system whereby voice to text recognition of conversations across telephone lines could allow a deaf person to be contacted by someone from an ordinary telephone. The functions of the prototype demonstrator are described hereafter.

In the second half of 1996, a prototype demonstrator was developed with just the basic necessary functions required and was presented to the users, stimulating their interest and providing a more precise feed back on their needs. The prototype made use of IBM VoiceType packages in Italian and English languages, with a piece of software developed by FBL for a more user-friendly presentation on the screen. Both IBM VoiceType 3.0 and Dragon Dictate were installed and tested and VoiceType was selected due to its user-friendliness and accuracy both before and after training. The importance of having high degrees of accuracy in both these situations lies in the fact that, whilst you need accurate recognition from a system to be used by a specific speaker in subtitling, you also need a software package that can recognise almost anyone's voice on the telephone.

In July 1997 we performed some tests with the new beta release of continuous speech recognition for the radiologists and in December 1997 we started the tests on the new releases of both IBM and Dragon continuous speech recognition. We improved the prototype demonstrators for the different applications foreseen in the *VOICE Project*. The introduction of *continuous* dictation greatly increased the effectiveness and potentials of voice to text recognition, so that in the first quarter of

1998 we could hold several presentations of the prototype demonstrator to Associations of the hearing impaired and schools. The suggestions received by them are helping us to bring the system nearer to the user requirements.

Other than the specific equipment related to voice to text recognition, the *VOICE Laboratory* hosts the *VOICE Project*'s Web site, which is accessible at all times across the Internet. The aim of this *VOICE Forum* is that of overcoming the communication problems of the deaf, demonstrating the usefulness of the Internet as a very appropriate means of contacting others and gathering information. A *VOICE Discussion Forum* and a *VOICE Chat Line* will also be provided, as a means of collecting and spreading information on the on going activities of the Project and more generally on voice to text recognition developments, facilities in tele-education programmes and technical aids for the deaf.

3.1. Subtitling conversations, school lessons and conferences

The text, generated by voice to text recognition commercial packages, may be of help for a deaf person during a normal conversation. But it is displayed in complicated windows and organised more in the style of a word processor rather than that of a subtitling programme. Following comments and suggestions from JRC-ISIS and users, FBL developed a piece of software that makes this more basic use of voice to text recognition user-friendlier. This software puts the generated text on a screen in various dimensions and colours, which can be set by the user, and only displays the part of the output from the recognition system which is of interest. This function of the demonstrator turns a voice to text recognition package into a subtitling system, i.e. with a certain number of lines, with a certain length and in a certain style.

A first version of this program had been created in the second semester of 1996 and demonstrated to the users in the first semester of 1997. At that time the system was based around the idea that it was still necessary that a *dictator interpreter* would dictate the text. This was foreseen with the aim of starting the activities as soon as possible, by using a prototype for gaining experience and developing awareness of the users, without waiting for new improved releases of the commercial packages. We used the voice to text package IBM VoiceType 3.0 that also possessed a function, called *VoiceType Direct*, allowing the user to select a text window into which he/she wished to dictate. This allowed to overcome some obstacles in integrating the programs.

The operating schema at that time was the following. The speaker, or in some cases a *dictator interpreter*, spoke into the microphone headset attached to the PC. The voice to text recognition package converted the spoken message into text. This text was displayed on the screen of the PC and also converted into convenient lines of subtitles that were passed via a network to a second PC. Here the subtitles-files were loaded at intervals and displayed on a black screen. The signal from this second PC to its monitor was passed through a piece of video overlaying hardware. This superimposed all the information (i.e. the subtitles) onto a video source (in most cases the image of the speaker taken by a video camera), converting the black subtitle background into transparent. The final result was therefore a composite video signal of the subtitled video source that could be viewed on a television set or recorded on a video-recorder.

This prototype demonstrator has been working at JRC-ISIS since February 1997 and has been installed in April 1997 at ALFA in Milan, for testing and use at the Association's meetings. After discussions with the users, tests and validation of the prototype, several improvements have been developed. The acquired experience helped us in further tests on the new releases of IBM and Dragon packages recognising continuous speech. This *VOICE Laboratory* is being enlarged with new releases of the voice to text packages as well as the development of new pieces of software, new PC interfaces and video signal mixing systems.

As since December 1997 (when submitting the first draft of the present paper), the working station configuration of the prototype demonstrator is based on a PC Pentium 200 MMX, with 64 MB Ram, Cd-Rom drive, audio card Creative SoundBlaster AWE 64, monitor 17". The software is Windows 95 operating system, Dragon Naturally Speaking or IBM Via Voice and additional software developed by FBL. In order to take the input of a video camera and send the final output to a video-recorder, a video card Matrox 4 MB with Rainbow Runner has also to be installed. The Matrox card gives the possibility of taking the input of a video-camera, while the processor Rainbow Ruimer, used in addition to the Matrox card, helps the CPU in managing better the screen.

The PC inputs are images and sounds. The images taken by the video camera are displayed on the screen, using the video card internal processor. The sound, i.e. the voice of the speaker, is acquired by the sound card and then analysed by the voice to text recognition package. The output is sent to the developed application program, which provides to manage the number of rows: either fixed upon specific requirements, or defined in automatic mode, controlled by the speaker's pauses. The text is displayed in a textbox, on a coloured background, at the bottom of the screen and is organised so that words are never divided between two lines. The user has the options of altering the font and size of the recognised text, the number of characters of text on each line and the number of lines of subtitles, as well as the colour of the background immediately around the subtitles. The generated text can also be saved and filed for future reference or use, as printing reports of conferences and minutes of meetings. This aspect, developed for the needs of the deaf, is of particular interest also for the hearing persons.

The complete demonstrator set uses a video camera and a wall projector, which are useful for conferences or television broadcasts. In the classroom or at home the system may be used without this additional equipment and the Matrox and Rainbow Raimer cards may be not installed. We have also performed tests on the use of a portable PC, which gives good performances, provided that a Creative SoundBlaster card is installed. We have used wireless microphones too, taking some additional care in setting the signal's input level.

3.2. Voice and pauses handling

The vocabularies included in the voice recognition commercial packages are of a general-purpose type, but easy to personalise. The voice packages are partially voice independent, but, in order to perform higher accuracy, it is necessary for the speaker to train the voice package (this operation requires an hour and is made once). Then it is important to check the dictionary with the text that will be dictated more frequently. The voice package detects the words not included in the dictionary and asks the speaker to type and dictate them, for adding them into the dictionary. As an alternative, the system may process the text (if already available on a file) in batch mode, discover all the new words and ask the user to dictate them, for adding them into the dictionary. Then the user has to train himself to manage the product, in order to balance the pauses as to handle short sentences and avoid breaking in the speech where not necessary, so as to get the best results.

We have often discussed with the users a particularly important aspect. Since lip-reading lets some uncertainty in interpreting the words, and also the speech recognition system may produce subtitling lines with some errors, the combination of both could help the users to get as near as possible to the originally spoken text. When our speech is addressed to an audience with a large number of deaf participants, or when they are just a few but we consider as a priority that they should get the best comprehension of what is being said, we use the prototype just as any other working tool.

So, we do not mind of a nice presentation effect, but we pronounce short phrases (for instance 3 to 6 words) that the participants may lip-read on the image of our face taken by a video camera and projected on the wall screen. Then we make a short pause (greater than 250 msec). The developed application program recognises this pause as a command of completing the recognition process of what has been said (this will take approximately an other 250 msec) and shows the generated text as subtitling lines on the wall screen. After having read the lines, we may, if necessary, repeat just one or two words that have not been recognised correctly. Otherwise we may continue the speech or add some details, if we consider it useful to repeat a word or use a synonym. We may do so, either because some words have not been recognised by the system, or because the audience seems not familiar with some particular words. If necessary, we may also type some words on the keyboard.

3.3. Subtitling television transmissions

This part of the demonstrator's development would allow to subtitle a video source using voice to text recognition. It concentrates on testing various ways for displaying the subtitles and video signals. We developed a prototype and since March 1997 we could input a signal from a video-recorder (or television aerial) into the PC, visualise the image on the screen and create subtitles using voice recognition software. The accuracy of the recognition directly from the signal was insufficient, but could be improved through the use of a *dictating interpreter*. Some further tests are foreseen and we are trying to get experience also by the use of a digital voice recorder as input to the recognition software.

Since subtitling of television transmissions is the result of a manual preparation of files to be transmitted in *Teletext* format, most of the subtitled transmissions are films. Subtitling of the news and of live programs, even those addressed to the deaf, is rarely performed. Voice to text recognition package could help in subtitling the films, speeding up the subtitling operations and probably reducing costs. This might allow to have more subtitled films transmitted by television broadcasters. Nevertheless this will not change dramatically the fhistrating isolation felt by the deaf community, who is looking for subtitling of live programs and of the news. For this more important aspect, the broadcasting companies might use voice recognition systems and produce good results. But the broadcasters should accept the risk of limited accuracy of recognition of the speakers' voice or should bear the cost of training the speakers to the use of the new systems and to build and update personalised new dietionaries.

An other approach could be eonsidered too. Instead of using the *Teletext* technology for broadcasting the generated subtitling lines, for specific transmissions the subtitling lines may be made available through the Internet. Subtitles lines across the Internet will still have a high refresh rate and the Web pages will only contain a few images. Uses for such a program could also be foreseen for the subtitling of radio broadcasts. The subtitles do not necessarily have to be created by broadcasting companies themselves. Independent members of the public, with the correct equipment and programs, could listen to the radio or television, summarise what is being said into a microphone and the subtitles will be broadeast world-wide over the Internet.

3.4. On the telephone line

This function of the demonstrator involves tests on a laboratory prototype of a computer-driven telephone, converting voiee to text, where a person ean speak down the phone line, the message will be passed into a PC at the deaf person's end and the words will be visualised on the deaf person's screen. In this situation *only* the deaf person would need the appropriate equipment, while the other speaker may use *any* ordinary phone.

As for the subtitling application, the eonfiguration for the telephone application is based on a PC Pentium 200 MMX, with 64 Mb Ram, Cd-Rom drive, audio card Creative Sound Blaster AWE 64, monitor 17". The video camera and the Matrox card are not necessary in this case, while a telephone line and a telephone set should of course be available. A filter has to be foreseen for decreasing the noise in the line and providing a better input signal. The software is Windows 95 operating system, Dragon Naturally Speaking or IBM Via Voice (which are used for voice recognition), Dragon Dictate for Windows V. 3.0 (which is used for managing the PC menus) and additional software developed by FBL.

The application will also include a text to speech system developed by MasterSoft to allow the deaf person to reply (should he/she have difficulties in speaking), which may also be useful in providing the person at the *dictating* end with feedback on whether what was said has been recognised correctly. It is worth noticing that when a hearing person is speaking at one end of a telephone line, and the text is showed at the deaf person's PC screen at the other end, the first user is *blind* with regards to the screen. Feedback of the recognised text is needed, just as a blind person needs text to speech synthesis to read a document from a PC. Taking into consideration this aspect, working for the deaf will also be beneficial to the blind.

The development of the prototype will, as first steps, improve the filters on the telephone line, make an qjplication program to manage the conversation between the workstation and people calling it and write the message on the screen. Experiments will also be carried out to see whether reducing the number of words in the software's vocabulary will improve the recognition accuracy.

The technology behind recent releases of voice to text recognition software will reduce the problems caused by the background noise present on telephone signals. The recognition engines of previous releases worked on dictated words being disjointed. Noise levels of signals taken from a phone line were so high during these necessary pauses between words that the quality of recognition was very poor. The latest releases from both the IBM and Dragon Systems use continuous dictation and no longer rely on periods of silence to separate words. Tests carried out recently (October 1997) using a beta release of the IBM Med-Speak software for radiologists have given encouraging results. Due to the limitations of bandwidth, as well as the noise on telephone lines, tests will be carried out on the signals from various lines (ISDN, GSM).

3.5. Costs

For the final user the cost of each demonstrator may be considerably lower than expected since some voice to text recognition commercial package is presently being sold at about 100 ECU (10% of the original price). This package contains an active microphone, whose individual cost is more than half that of the complete package.

The approximate cost for the basic structure on which the applications are based (a Pentium PC with a SoundBlaster16 Sound-card, a Cd-Rom drive) is 1500 ECU, or 2500 for users demanding particular applications (video mixing, etc.). This price nevertheless also includes a fully operational PC that can be used in many other useful ways. A great advantage is also the fact that the system is not dependent on any particular company or software release. The equipment at the application sites (FBL laboratory and JRC *VOICE Laboratory)* is however slightly more expensive. The added costs are due to increased network capabilities and additional pieces of hardware and software to be used for everyday work as well as for comparisons, testing and demonstration purposes.

4. Results and final goal

Continuing on a well tested procedure, each step of the prototype is developed by FBL with the tight collaboration of JRC-ISIS *VOICE Laboratory*. It has been presented to the users and discussed and tested with them. Until now, when the opportunity has arisen, the demonstrator has been presented to conferences or general assembly of the Associations of the deaf, for giving awareness on the possibilities of the system. In the first quarter of 1998, the system has been presented to some schools, where it will be tested in real situations of use, with several constraints and specific difficulties. It will be used for subtitling school lessons (for the benefit of the deaf students), as well as for the visualisation of the dialogue pronounced during the foreign language lessons (for the benefit of the hearing students) or of the lessons on the host country's language (for the benefit of any user, particularly the immigrated). Some tests have been also foreseen for subtitling university lessons and printing summaries.

In the next months, we will use the prototype demonstrator for subtitling the speech of some speakers in conferences, this being an important opportunity of testing and validating the demonstrator in different operating conditions. During the ICCHP-98 Conference we will present the system to the users in an international environment and test it in English, Italian and German languages, in real operating conditions. The users will be encouraged to give their comments and suggestions in a *VOICE* Workshop and in the *VOICE Special Interest User Group*, which will hold its first meeting during ICCHP-98. The suggestions will help in improving the prototype in order to show other functions to the hearing impaired, the producers of speech recognition systems and the services providers.

This development is foreseen with the aim of discussing of new possible functions as soon as possible, gaining experience and developing awareness of the users, without waiting for new improved releases of the commercial packages. But, as it has been stated in a previous paper (*The VOICE Project - Part 1*) presented to ICCHP-98, the final aim of the prototype demonstrator is *not*, as in many other projects, to increase its size and its performances. The aim is just ... to disappear! We hope that during the two years of the Project's life, we will be able to help the hearing impaired to convince the producers of speech recognition systems of the interest for them to include some of the proposed basic functions into the new releases of their standard products, for the benefit of any *user.*

Session II

Blindness: Access to Documents

A GUI-based OPAC for the Austrian catalogue for alternative format books

Franz Burger, Gerold Wagner
Modellversuch „Informatik für Blinde“
Institut für Angewandte Informatik, Universität Linz
{franz|gerold}@mvblind.uni-linz.ac.at

Abstract

Searching in library catalogue systems is often impossible for visually impaired readers due to the lack of accessible on-line public access catalogues (OPAC). The reason for this is that the applications are not designed for access with adaptive technology as used by blind and partially sighted users.

Part of the Austrian workpackage in the TESTLAB project was the development of an OPAC which is based on a graphical user interface (GUI). The application has been designed to allow a maximum of accessibility.

A set of user interface design principles has been derived from the actual implementation of the TESTLAB OPAC. Usability tests have shown that careful design of GUI-based OPAC may open the library world to visually impaired readers by providing an easy-to-use and efficient search tool.

1. Background

The Educational Endeavour "Computer Science for the Blind", Johannes Kepler University of Linz, Austria, is currently partner in the European Union funded TESTLAB project. The main aim of TESTLAB is to use information and telecommunication technology to give visually impaired readers as near as possible the same level of access to catalogues and documents as sighted users. TESTLAB will establish a series of practical trials in public and academic libraries whereby blind and visually impaired readers can gain access to catalogues and documents in forms which they can read.

The University of Linz, in co-operation with the University Library of Graz, is developing a library catalogue for literature in alternative formats. This library catalogue will enable visually impaired readers to search for books in alternative format (Braille, large print, audio cassette, digital) available in Austria by means of a computer.

Blind computer users may gain access to information systems via adaptive equipment like speech synthesisers or transitory braille displays. This devices translate the information shown on the screen into a perceptible format for them. Partially sighted readers may use screen magnifiers which enable to magnify the display as well as to chance colour and contrast appropriately. [6]

The library catalogue developed in TESTLAB not only enables visually impaired users to search for literature, but also allows university libraries to co-ordinate the transcription of study material more effectively.

An extended version of the catalogue includes records from German special library catalogues which accounts for a total of more than 58,000 titles. As an example, the catalogue also covers bibliographic data on the "digital books" from project Gutenberg-DE (http://gutenberg.aol.de) available on the Internet. Consequently, visually impaired users have instant access to this digital library and may read books by means of a computer.

For more information about the TESTLAB-project visit the web pages at http://www.svb.nl/ or http://www.mvblind.uni-linz.ac.at/.

2. The Austrian TESTLAB catalogue system

All the catalogue data originating from the various special libraries and catalogue providers are available in digital format. However, the structure and completeness of the data provided differ a lot. The catalogue data are delivered in various formats (e.g. MS Access, dbase, structured text).

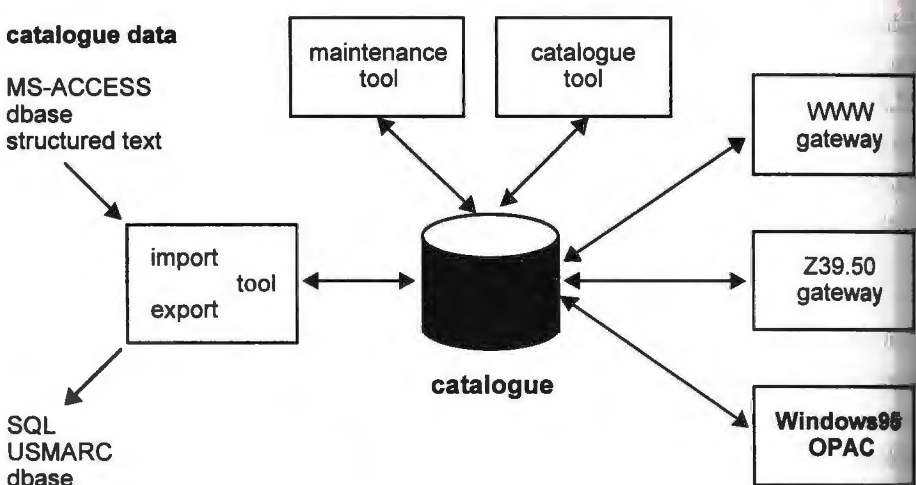


Figure 1 Library Catalogue System

The actual storage of catalogue data is provided by a commercial database management system (DBMS). The DBMS uses a relational data model. All data definition and manipulation is carried out with the Sructured Query Language (SQL). As a consequence, the whole database does not rely on a vendor specific product but can easily be transferred to other systems. This is particularly required for the provision of alternative search facilities (Win95 off-line OPAC and Internet-based search).

The import tool is responsible for transferring catalogue data into the central database system. The main task carried out by the *Import Tool* is the recognition of bibliographic records and proper insertion into the database. Since the data records are originating from various sources and do not conform to common cataloguing rules the import tool has to provide for basic data harmonisation.

The export of catalogue data either to another DBMS or to another format is provided by an *Export Tool*. This program allows the selection of the whole database or just a sub-set of records matching predefined criteria (e.g. library code, format). The range of record formats supported so far cover SQL, USMARC and the dbase format.

The maintenance of the catalogue by the system administrator can be done by a special *Maintenance Tool*. This program provides facilities for improving the integrity of the whole database. New bibliographic records may as well be added with this tool to the catalogue database. A subset of the facilities provided by the maintenance tool is available also in the *Catalogue Tool*. It will be mainly used by librarians to add new or edit existing records.

Visually impaired users will have various possibilities to query the catalogue. The primary OPAC application is an application based on a graphical user interface (Windows95). The application has been designed with respect to accessibility criteria for adaptive software and may be used for both, on-line searches on the central database and also as a standalone application searching a local database. The requirements and user needs for the query applications have been derived from user tests on OPAC access [1].

As the majority of visually impaired library patrons still prefers to use software with a text-oriented user interface [4] an alternative OPAC application developed and maintained by the University Library of Dortmund can be used for searching.

The range of search facilities for end-users will be completed by two alternatives accessible via the Internet:

1. A web-gateway will present a search form by way of any standard web-browser to the end-user and perform searches on the central database.

2. A Z39.50 search interface provides a standard way to query for bibliographic records using a standard search and retrieval protocol.

3. The Windows OPAC

The primary OPAC application has been developed for use with Microsoft Windows, version 3.1 or later. The full installation consist of the following software components:

- Microsoft Windows version 3.1 or later
- Gupta/Centura SQLWindows 5.0 runtime environment. This software is available for free.
- Database: To run the off-line version, Gupta/Centura SQLBase 5.2 has to be installed. The limited version that supports only databases not larger than 5 MB is freely available.
- Database customisations: The sql.ini and country.sql which contain customisations according to the German language have to be replaced by slightly modified versions.
- Adaptive software: If necessary, the adaptive software has to be installed.

The OPAC has so far been tested with the following software:

- Virgo 2.0 by Baum Products: This software and a small customisation for the SQLWindows objects are required. The customisation will be delivered together with the OPAC.
- JAWS 2.x by Henter-Joyce: This software can be used without special customisations.
- Ai Squared ZoomText.

3.1. Navigation

The navigation is a basic fundamental for exploring the user interface. While sighted users can get an impression of the functionality of a user interface within a few seconds, for blind users it causes much more efforts to get this overview impression [4]. Therefore, it is most important to optimise the possibilities of navigation. The OPAC has been designed especially for the support of the following navigation methods:

Navigation by thumb keys: This kind of navigation is supported by screen reader software and involves the braille display thumb keys for navigating through the user interface. Efforts have been made to improve the comfort of thumb key usage. More details on this will be described in the section "Optimisation for visually handicapped users".

Navigation by keyboard using the tab key: Most GUI applications designed for sighted users can easily be navigated by use of the mouse. Since mouse usage is generally impossible for blind users it is recommended to offer comfortable keyboard navigation. This kind of navigation is implemented in Microsoft Windows by use of the tab key: The tab key is used to move the cursor from one user interface element to the next. This successor is not necessarily positioned spatially near to its predecessor since the sequence of elements is defined by the application developer. The sequence of elements may be unimportant for the sighted user because he can get information about the logical context of the graphical items out of their locations on the screen. For the visually impaired user it may cause serious confusion to jump from one element to a non-related successor.

Keyboard accelerators (short cuts): Keyboard accelerators can be very helpful for jumping to a well-known element of the user interface or for triggering an action. Therefore keyboard accelerators are defined for almost every element.

3.2. Customisation

If the user wants, he can customise some features of the OPAC. All the customisation settings are stored in the database so that the user can define them only once and they will be loaded automatically at each start of the OPAC. Currently four different OPAC features can be customised:

Complexity of search form. The user can choose one of three levels of complexity.

- The simplest level provides a case insensitive search by the attributes "author" or "title". For beginners this may be enough, so they should not be confused by additional screen elements they would never like to use.
- The medium level provides the additional search criteria "library", "media type" and "subject".
- The complex level provides additional search fields for "publisher", "keyword", "year from" and "year to" as well as the possibility to choose between case sensitive and case insensitive queries. The Boolean operators „AND" and "OR" are as well supported for query formulation.

Web browser. Internet on-line documents can be instantly read by the use of web browser. If the user specifies which web browser is installed on his PC, he will be able to load an on-line document referenced in a database record by simply typing "Alt+w" or hitting the corresponding push button.

Font face and size. Partially sighted users can adjust the font face and size independently from the Windows default settings according to his needs.

Colours. This is also a feature for partially sighted users who can set up their own colour settings and store them in the database. Background and text colour can be adjusted independently to get the preferred screen display.

3.3. Queries

As this is the nature of OPACs, the most important feature is to perform queries on the records stored in the database. A query can be performed in a simple way:

Immediately after the start of the OPAC the caret is set to the "author" input field. The user may type the last name of the author. If (s)he wants, (s)he can also (or instead of specifying the author) type parts of the title into the "title" input field, which the user can reach by hitting the Tab key or Alt+T. Once the user has specified author and/or title, he can perform the query by hitting Alt+S or activating the search button. A beep signals that the query has been executed and the focus is automatically set to the result list of titles. This result list displays the combinations of authors and titles that match the query. The user can navigate in this list of items using the arrow keys, the braille bar thumb keys or the mouse. Once having selected an item, the user can activate the details view by hitting the Enter key, Alt+d, double clicking or activating the "detailed-informations" push button. The details dialog displays a list of all the copies that are available for the selected titles and additional information like media type, library, location, signature, URL etc.

3.4. Display on-line documents

If the user has previously specified his web browser in the customisations dialog he may launch the web browser starting from the details dialog: If the media type of the selected copy is "on-line", the user can view and download this document with his web browser by hitting Alt+w or the "www-browser" push button. In this case the web browser is started with the URL of the document which means that the desired on-line-document is automatically loaded into the web browser.

3.5. Save query results

The user may save both result lists, the list of titles and also the list of copies in the details dialog, to a local text file.

4. GUI Optimisation for visually handicapped users

The optimisation in respect to the needs of visually handicapped users has turned out to be quite sophisticated. The general direction of our efforts was to provide a clear user interface to braille and speech users on the one hand and to allow customisations for partially sighted users on the other hand. The general endeavour was that we tried to combine the advantages that Windows brings for visually handicapped persons with the possible simplicity known from DOS based user interfaces.

The OPAC has been tested with *but not specifically designed for* two different screen readers, namely "Virgo 2.0" offered by Baum Products and "JAWS 2.x" offered by Henter-Joyce as well as with the magnification software "ZoomText" by Ai Squared. In the following sections, general features on the one hand and selected types of graphical elements on the other hand will be described.

4.1. General features

In order to provide a standardised, clear and customisable user interface some principal guidelines have been followed;

- Each of the elements of the user interface has its own keyboard accelerator
- The user can choose one out of three levels of complexity according to his needs or interests, starting from the "simple" option that only provides search just for author and title and reaching to the "complex" options that provides all query options that are possible.
- The OPAC has been designed with regard to non-advanced users: The user does not need to be an expert in using adaptive software or navigating with braille bars. The only preconditions for using the OPAC are that the user must be able to type on the keyboard and to use the tab key to jump from input field to input field (and this only if he does not want to use the short cuts).
- Provide an intuitive interface that permits direct, informed interaction [5].

4.2. Colours & Fonts

Although colours and fonts are unimportant for blind users, special colour or font settings can be very helpful for partially sighted users. The OPAC supports colour and font customisation in two directions:

- The first, simple but useful feature is that the OPAC uses the colour and font settings of the Windows environment, if the user wants this.
- A second possibility for the user is to specify his or her own colour and font settings for the OPAC and to save these settings for re-use in later sessions

4.3. Descriptive Texts

The descriptive text (label) is, at least for visually impaired persons, the most important help for users to find out one's way in user interfaces. So the following features have been regarded during the design of the screens:

Location. The descriptive texts are always placed immediately left of or above the corresponding items. Short distances between the labels and the editable items are of extreme importance for some adaptive software, especially screen readers.

Short cuts. The short cuts that the user can use for accessing specific screen elements are displayed in the labels. This is done in a way that adaptive software is able to identify and display the shortcut. Each of the GUI elements has its own short cut.

No segmentation of the user interface by the use of special characters. Separations of distinct parts of the user interface by the use of textual elements have been avoided since they often seem to confuse blind users. Wherever separation has been useful, it has been implemented by the use of graphical elements that do not confuse screen reading programs.

4.4. Edit Fields

So called edit or data fields are used for input of most of the query attributes. Normally, edit fields do not cause problems in terms of accessibility if the programmer uses left bound texts and does not leave too much space between the labels or the left window border and the edit field.

However, there is a standard Windows feature that confuses braille bar users or at least complicates the usage: Whenever there is a text contained in an edit field and the user sets the focus to this field using the tab key, the contained text is highlighted and the caret is set to the position after the last character. If the text in the edit field is longer than the Braille display, the braille bar user can read only the end of the text on his display. To read the beginning of the text in the edit field and the label extra navigation is required. To avoid this unfavourable behaviour the Windows standard functionality has been overwritten: If the user tabs to the edit field, the caret is set to the first character which allows the braille bar user to read both the label and the beginning of the text.

Additional changes were necessary since one of the screen reader software packages has problems with empty edit fields. To work around this bug of the adaptive software, empty data fields are by default filled with a blank character. This leads to reliable recognition of "empty" data fields.

It has come out that it is useful to use edit fields not only for input, but also for output. Many developers use descriptive text as objects for output. Since these objects cannot get the cursor focus, users could fail to notice them. Therefore edit fields that can get the focus (although they cannot be edited) should be preferred.

4.5. Combo Boxes

Combo boxes or drop down list boxes were originally created to save space on user screens on the one hand and to provide the same features as single selection list boxes on the other hand. Based on this definition the use of combo boxes would not be useful for the user interface design for blind users. But the possibility to combine the input feature of edit fields with the list box features makes sense in the OPAC user interface. According to the skill or the preferences of the user he can either type in the data or select one entry out of a list or combine these two ways by typing the beginning of the entry and hitting the arrow-down key once. One should be aware that user interfaces designed for sighted users would most probably *not* use *editable* combo boxes, because sighted users who may find it simple to select one of a *non editable* combo box entries by use of the mouse. Additionally, the standard features of the combo boxes have been extended for the use by blind users in another way. If the user has already selected the first entry and he presses the "up"-key, an error-beep occurs. The same behaviour is implemented when the user presses the "down"-key although he has already reached the last list entry.

4.6. Menus

Menus are mostly known by blind users from their experience with DOS applications. Not all DOS menus really support the use of braille bars and often the user has to search where the menu items can be found. In contrast, standard GUI menus are in general better accessible than DOS menus because the braille cursor follows the selected menu items. As menu items can be activated also with short cuts, they are in general a pleasant GUI element for visually handicapped users.

5. Experiences & Recommendations

Our experiences are that it is possible to create applications with GUI that are easy to use by users of adaptive software. The most important preconditions for the development of accessible applications are that the program developer is aware of the concepts of adaptive software and follows some rules that are mostly not specific to the usability by blind users, but more general:

Navigation

- Define a label and a keyboard accelerator for each element
- The logical order of the elements should be equivalent to their position on the screen
- Support multiple ways of navigation
- Use simple structured screens

User interface objects

- Use standard standard Windows objects: The user will already know them. Using them will avoid implementing custom elements which probably will cause troubles to adaptive software.
- Place related objects near to each other
- Overwrite or disable disturbing standard MS Windows features
- It may be necessary to add some useful features to work around bugs of adaptive software or to increase user's comfort

Standards & Customisations

- Use standard environment settings; A user who works with MS Windows usually has already changed the settings according to his favours, so why should you define deviating colours or fonts?
- Offer customisations where useful: different complexities of the user interface will be honoured not only by visually impaired users.
- Accessible user screens can be visually attractive. Be aware of the specifics needed by visually impaired users and you can develop both in one, an attractive *and* accessible user interface.

Acknowledgements

This work is being carried out in the context of the TESTLAB project (LB 4003) funded by DG XIII of the European Commission under its Telematics Applications Programme, Libraries Sector, and by the State Government of Upper Austria, Department for Social Affairs.

References

[1] BURGER, F., Report on OPAC Access by Visually Handicapped Users. Deliverable 4.1, TESTLAB project (December 1996). http://www.svb.nl/project/download/test-4-1.zip

[2] DEINES-JONES, C., Access to Library Internet Services for Patrons with Disabilities: Pragmatic Consideration for Developers. In: Information Technology and Disabilities, Volume 2, Number 4 (December 1995). http://www.rit.edu/~easi/itd/itdv02n4/article5.html

[3] EDWARDS, A. D. N., The Rise of the Graphical User Interface. In: Information Technology and Disabilities, Volume 2, Number 4 (December 1995). http://www.rit.edu/~easi/itd/itdv02n4/article3.html

[4] HILTON-CHALFEN, D., Information Technology, Campus Libraries, and Patrons with Disabilities: Emerging Issues and Access Strategies. In: EDUCOM Review (Nov/Dec 1993).

[5] HILDRETH, CH. R., The GUI OPAC: Approach with Caution. In: The Public-Access Computer Systems Review, Volume 6, Number 5 (1995). http://info.lib.uh.edu/pr/v6/n5/craw6n5.html

[6] LINDERS, P., VERBOOM, M. AND LEEMAN A., Report on the Inventory of Interfaces, User Needs and Existing Adaptations. Deliverable 6.1, TESTLAB project (October 1997). http://www.svb.nl/project/download/test-6-1.zip

[7] SHNEIDERMAN, B., Designing the User Interface: Strategies for Effective Human-Compu Interaction. 2nd ed., Reading, MA: Addison-Wesley (1992).

An electronic daily newspaper for blind readers

P. Graziani, L. Burzagli, E. Palchetti [1])

Abstract

We describe a service for blind people developed in Italy in the framework of the National Project on Telemedicine of the Ministry of University and Scientific and Technological Research. This service consists in a Web site containing a completely accessible version of one of the most diffused Italian daily newspapers. This electronic newspaper is automatically generated with materials provided by the publisher. Particular attention has been devoted to the navigation effectiveness inside the newspaper with the study of a suitable document structure and the development of some special reading tools.

1) National Research Council (CNR)
Inst. of Research on Electromagnetic Waves (IROE) "Nello Carrara" Florence (Italy)
Via Panciatichi, 64 – 50127 Florence (Italy)

1 Introduction

Text Blind persons have achieved in the last 15 years a large access to information sources, included printed paper, by means of the personal computer equipped with a scanner with OCR programs, speech synthesis or Braille display. Nevertheless, the daily newspaper remains very difficult to be read with these technical systems because of the poor print quality, the complex graphical structure and the format generally not compatible with the size of the scanner. A number of projects have tried to find solutions to this problem. For example, in the 80's a service of nightly radio distribution of an encoded version of a newspaper was organised in Sweden for blind users equipped with a special radio-receiver connected to their computer [3]. In Italy, since 1990, a large selection of articles of a daily newspaper (La Stampa of Turin) has been distributed to blind persons via Telesoftware pages, the binary pages associated to the Teletext service (called Televideo) of RAI, the Italian Broadcasting Corporation. These pages can be received with a decoder card for p.c. and saved on the hard disk as ASCII file, to be read like any other similar file.

A TIDE project, "Communication and Access to information for Persons with Special needs" (CAPS, 1991-94), has defined an European standard format for the interchange of electronic newspapers, facing problems of different classes of "print disabled" [2]. This format is based on SGML (Standard Generalized Mark-up Language, ISO 8879) and consists of a CAPS News DTD (Document Type Description). A journal in this format can be read with a special CAPS reader program.

Many daily newspapers and other periodicals are now available in HTML format in a number of Web sites. Generally this represents a very important source of information for people unable to access the normal printed form of these publications. In principle, also blind people can benefit of this new possibility, thanks to the availability of special browsers or screen readers for DOS and Windows making normal browsers accessible for the blind through a speech synthesizer or a Braille display.

Unfortunately, in many cases, sites of newspapers present more or less severe accessibility problems for blind readers, so that these potentially powerful new information sources can be either used only by particularly expert blind users able to deal with a complex graphical user interface, or cannot be used at all.

We can observe that, even if many organisations are studying the general problem of Web sites accessibility and guidelines for the design of accessible HTML documents are available, most designers ignore these problems with the result of the creation of artificial barriers which could be easily avoided by paying more attention to these aspects.

Taking into account this reality, the National Project of Telemedicine of the Italian Ministry of University and Scientific and Technological Research (MURST) has supported a development activity to provide a service of distribution through a Web site of at least one daily newspaper completely and easily accessible for blind readers. The MURST

Telemedicine project consists of 8 different sections, one of which is devoted to the development of support systems for disabled people, with a consortium of different institutions, with CSELT of Turin as prime contractor. The development of an accessible electronic newspaper for blind persons is one task of this project section and it is carried out by IROE. Another partner is the Italian Blind Union undertaking the task of analysis of user needs and evaluation of the results.

2 The structure of a newspaper

A daily newspaper does not consist of a single text to be read from the beginning up to the end. A sighted person generally takes a glance at a newspaper in order to choose what is really interesting and what not. In many cases, the text of an article is skimmed through rather than really read, that is the reader skips portions of paragraphs and sentences, in order to approximately realise which are the basic concepts expressed by the writer.

A blind person who reads an electronic version of the newspaper, finds many difficulties in the exploration of the contents and in browsing an article. A sequential transducer, such as a speech synthesizer, or a short line of transitory Braille, does not allow the overview of a document.

This means that a newspaper for blind readers must be organised in such a way to facilitate its exploration and a quick reading of articles.

The hypertext structure solves this problem, at least for what concerns the overview and the effective contents exploration. In our project, a set of HTML pages has been organised into a tree structure to represent the different levels of the logical classification of the different components of the newspaper.

Incidentally, it can be observed that the tree appears as a very good metaphor for a blind person since it allows an organisation of the exploration following several sequential paths, according to the choices of the reader, well compatible with the features of the alternative display, with an easy mental reconstruction of the newspaper structure.

A double tree structure has been chosen to present the contents of our electronic newspaper according two different criteria: a logical classification of titles according to subjects and a physical location organised into pages. The first key of access allows a quick exploration of lists of titles with a common subject, regardless of the position of these articles inside the printed version of the newspaper; on the other hand the second key of access brings the blind reader into the same condition of a sighted person dealing with the printed edition, where the reader browses the journal page by page, to descry their contents and the articles distribution and location.

Both sets of title lists (subjects or pages) consist of titles in short format, without head-lines and sub-titles, in order to make the blind reader able to quickly explore each list and to select an article. Each title represents a link to a short page detailing the selected article. First of all, each component of the complete title is presented with the correspond label: "Head-line", "Title" and "Sub-title". In addition, the page number and the position of the article in the page is specified with the name of the author and the file size in bytes. In particular, this last piece of information allows the reader to realise how large is the article inside the page. In fact, in the articles list each title is written with the same font so that it is not possible to realise the extent of the corresponding article in the printed version. This is an import difference between the reading of the two forms of newspaper.

Finally, three buttons give access to three different reading modalities. The main one is the integral reading, that is the sequential presentation of the complete text of the article. The second and third ones represent two different forms of quick reading consisting of pages generated "on the fly" with a reduced presentation of the text: only the first portion of each paragraph or each sentence, respectively.

These forms of reading, based on sentences cut after a pre-defined number of words, try to reproduce the browsing of an article that a sighted reader performs by jumping from the beginning of a paragraph to the next one, befo concluding the reading of each paragraph.

The two implemented levels of quick reading (paragraphs and sentences) allow to choose the most suitable one for each case, according to the structure and the style of each article. In the case of large paragraphs, jumping from one paragraph to another can be too much to allow a sufficient understanding of the article. In this case, the presentation of the beginning of each sentence can be more suitable to have a general idea of the contents.

As default, the reduced text form consists of paragraphs or sentences cut after 10 words but the user can choose ano length by means of a small form associated to the quick reading buttons.

A preliminary test demonstrated the effectiveness of these forms of quick reading.

3 Implementation of the electronic newspaper

The chosen daily newspaper is La Stampa of Turin, one of the most common and known newspapers in Italy. Th company publisher of this newspaper has offered the necessary cooperation by making available the files of artic written in their proprietary code, and all useful information to interpret such an encoding. The material of the newsp is made available by the publisher free of charge to the MURST-Telemedicine project for blind users.

Every night, a compressed file, containing all articles of the newspaper, is sent to IROE via e-mail by the Turin editori office. An automatic procedure extracts the compressed file attached to the e-mail message, explodes it into a work directory of the server where all articles are converted, in several steps, up to the final version of the electr newspaper consisting in the structured set of HTML pages described above. These pages are immediately available at the URL of the server for blind readers, generally starting from 1 or 2 a.m.

The first stage of the conversion necessarily consists of the interpretation of the proprietary code of La Stampa but all the re-organisation of the newspaper is designed in such a way to allow the same process to be applied to other newspapers which could be available for this service in a near future. We adopted an intermediary format which is the basis of the whole generation of the final version structure. If another newspaper is available, written with another code, only the first conversion, from this code to the intermediary one, will require the development of a new converter, while the main part of the process will remain the same.

The intermediary format chosen for our newspaper is the CAPS News DTD [2], mentioned in the introduction, which appears well designed to organise all the newspaper components.

All software packages are installed on a Sun Ultra 1 workstation. This machine has been equipped with an Apaches httpd server for the implementation of the www site, an "anonymous" FTP server for the down load facility, a Majordomo listserver for creation of mailing lists and the Webglimpse search engine.

The cascade of programs necessary to process the newspaper materials, from the e-mail message up to the HTML pages set, consists of several software modules most of which have been developed in the frame of the project while other ones are freeware programs available in the network, used to perform particular functions such as Mime decoding and decompression of the file extracted by the e-mail message.

Moreover, we installed on the same machine the CERN httpd server in order to implement the special proxy server W3-Access for the blind, developed by the Institut fuer Informationssysteme, ETH Zurich, to make available its powerful on line filter, allowing to improve the access possibilities to any Web site.

4 Distribution of the electronic newspaper

Blind readers can use the electronic daily newspaper in different ways. First of all they can read the newspaper on line on the server, via an Internet address (http://sunmurst.iroe.fi.cnr.it or http://etabeta.iroe.fi.cnr.it). This allows the user to explore the journal contents, following the classification by subjects or pages, to browse some articles, to search for some strings inside the whole hypertext. The latter function is performed by means of a search engine installed on the server and represents an access key to the newspaper following the user's personal criteria.

Because of the cost of the on line connection, the reading of large articles with this modality can be too expensive. For this reason, a compressed file is available to be down loaded and exploded on the user's hard disk for an off line reading.

This compressed file contains all HTML pages of the newspaper, including tables of contents, so that users can read the newspaper by using in local mode the same browser used for the Internet connection. Of course, in this case some server facilities, such as search engine and quick reading commands, are not available.

A third possibility is represented by a textual version of the newspaper, also available in a compressed file form organised into a structured set of pages, like the HTML version, but in a form compatible with the DOC program, a textual browser developed by Giuliano Artico of University of Padua. This program is well known and widely used in Italy by many blind computer users to read ASCII files in general.

Incidentally, we can observe that the ASCII version of the newspaper can be read by means of a DOS program running on any machine. As an example, a blind commuter declared that he usually downloads the compressed ASCII file from the server, early in the morning, he transfers it into an Audiobraille (a small portable device provided with DOS, Perkins keyboard and speech output), and then he reads the newspaper on the train!
A listserver is installed on the same machine of the httpd server so that a distribution via e-mail of the compressed versions of the newspaper can be easily organised, if users request it.

5 First result of the experimentation

A preliminary testing was performed in order to verify the different aspects of compatibility of the Web site with the variety of browsers used by blind users.

HTML pages of the newspapers have been designed in full accordance with the accessibility guidelines of Trace R and D Center, University of Wisconsin, Madison [5],[6], and validated with "Bobby", the tools made available by the CAST (Center for Applied Special Technology).

The complete accessibility of the newspaper and all other pages of the server has been verified by users, regardless the browser used.

A separate test was done with a controlled access through user name and password given by means of a small form. In this case, people using Minuet and Nettamer (special DOS browsers for the blind) were not able to access the newspaper while no problem was found by people using normal browsers, such as Lynx, Netscape or MS Internet Explorer, in combination with a screen reader for DOS or Windows. This means that, in the case that the newspaper publisher requests this controlled access, many users will have to change their browser.

Concerning the newspaper organisation, all users found the double tree structure of HTML pages effective and easy to explore.

Only a few users tested and generally appreciated the quick reading facilities.

A more general evaluation will be conduced in cooperation with the Italian Blind Union. This evaluation will concern the whole development of the project, including performances of a special user terminal and of an of electronic book encoding system carried out by other partners.

6 Conclusion

The electronic daily newspaper developed in the framework of the MURST-Telemedicine project represent an example of Web site, which was designed taking into particular care, not only general accessibility problems, but also the needs of navigation effectiveness inside an hypertext by means of the typical technical aids used by blind persons and the particular nature of the stored information.

This is important in order to encourage even non experienced blind people to access these new and powerful information sources.

This paper described just the development of the electronic newspaper, but other important aspects of the more general subject of an easier access to other newspapers available in Internet have been taken into consideration by the mentioned project. For example, on the same server, the user finds important information about special proxy servers which can be used in combination with a browser to improve the accessibility, such as those developed by the Institut fuer Informationssysteme, ETH Zurich, dedicated to the filtering of HTML pages [4], and by Adobe Acrobat Products, to convert PDF into HTML pages [1].

Partners of the same project have carried out other important developments but this may not be described in this paper.

7 References

[1] Adobe, (1997) Helping people with visual disabilities access electronic information, http://www.adobe.com/prodindex/acrobat/access.html.

[2] CAPS, (1992) Communications and Access to Information for Persons with Special Needs Task 2.1 EIF Specification Report Deliverable 2.1 October 1992.

[3] Lindstrom, J-I, (1989) Teleinformatics and disability in Proceedings of Conference on Computers for handicapped persons , AUSTRIAN COMPUTER SOCIETY (OCG), Wien, 39 401.

[4] Perrochon, A. and Kennel, A. (1996) WAB: World Wide Web Access for Blind and Visually Impaired Computer Users, http://www.inf.ethz.ch/department/IS/ea/blinds/.

[5] Vanderheiden, G. C. (1998) Unified Web Site Accessibility Guidelines, January 20, 1998, http://trace.wisc.edu/docs/html_guidelines/htmlgide.htm.

[6] WAI, (1998) WAI Accessibility Guidelines: Page Authoring W3C Working Draft 03-Feb-1998, http://www.w3.org/TR/WD-WAI-PAGEAUTH-0203.html.

EXPERIENCE WITH HYBRID BOOKS FOR BLIND PEOPLE IN ITALY AND HUNGARY

Paolo Graziani (graziani@iroe.iroe.fi.cnr.it),

Inst, of Research on Electromagnetic Waves

(IROE - CNR), Florence, Italy

Teréz Vaspöri (vaspori@sunserv.kfki.hu)

Laboratory of Speech Technology for Rehabilitation

(LSTR - RMKI), Budapest, Hungary

Abstract

CD drives and sound cards are becoming standard components of the personal computer, thus enabling the Copernicus DIGIBOOK project (coordinated by Prof. Jan Engelen from KU Leuven, Belgium) to introduce the so-called "hybrid book" concept. According to this concept, the structure of a book is stored together with its ASCII text and digitally recorded and compressed human speech reciting the book text. The hybrid book processing system and the browser were developed in Hungary by the Laboratory of Speech Technology for Rehabilitation (LSTR) led by Dr. András Arató. This article describes the results and conclusions of the first experiments performed in Italy (CNR-INROE, the Italian member of the Copernicus DIGIBOOK project) and in Hungary.

Introduction

The cost of the technology to record CD ROM has rapidly decreased over the last few years so that the production of such disks has become available also for products devoted to restricted computer user groups, such as blind people. Electronic libraries containing the books in pure ASCII text format have also been founded. In Italy, electronic books retrieved by scanner and OCR programs are developing, distributed on CD ROMs. In Hungary such electronic text books are distributed on diskettes for blind people. To create a hybrid book, a simple ASCII file is translated into a form that is slightly different from an HTML. This form contains both the structure and the text of the book. The concept and creation of the hybrid book is discussed in detail by A. Arató and L. Buday [1]. The HTML-like source and the human speech recordings form the input to the pre-compiler. The purpose of the pre-compiler is to create three separate files containing the structure, the text, and the voice of the human performer. The structure file has been created in order to enable fast and effective navigation and to minimize the CD-ROM search time during the DIGIBOOK reader program.

The Talking Book Service of the Italian Blind Union has already started to record talking books in digital form to store masters on CD ROM. They have expressed interest in the DIGIBOOK project and participated in this evaluation. The Hungarian Blind Union (HBU) is taking part in an ongoing national project to develop a special hardware device for hybrid books in Budapest. The studio of the HBU will produce hybrid books for Hungarian blind users applying this technology.

Appraisal of the hybrid book in Hungary

During the DIGIBOOK projects two Hungarian, one English and one Italian hybrid books were created and tested by using a specific hybrid book reading software. A questionnaire was formulated by the developers of the browser in connection with the reading software. Understanding the concept of the hybrid book and the use of its reading system were taken into consideration in the questionnaire. The Hungarian consortium member invited 12 blind users in four groups (3 people to each group) to the laboratory for testing the books and the program. An introduction was given to the testers and time was allotted for them to use the books. Upon completion, they were asked to fill out the questionnaire.

The result of this survey shows that all participants understood the structure and basic entities of the hybrid book. 11 people were able to start the browser and to read the books, one person needed help. 3 people judged the browser simple, logical and perfect, 9 people gave suggestions for solving same functions in another way. All 12 people stated that after some hours of training each service of the browser could be used easily and they would listen to such books with pleasure.

Description of the experiment in Italy

In Italy the chosen text was a tourist guide of Rome, which is suitable to show the possibility of navigation and random access to any part of the text. The Italian experiment showed more interesting results. On one hand, the experimental Italian Hybrid book has been developed by adopting the Hungarian method, in particular, the pre-compiler and the browser developed in Hungary was utilized. On the other hand, after self-made installation, users tried to read the book using different synthesizers other than the supported one (BraiLab).

Due to the structure of the browser, with the language dependent section written in a separate file, this program could be easily adapted to the Italian language. The browser was tailored to the Hungarian text to speech system (BraiLab), development of a small resident program was necessary to capture messages sent by the browser directly to the virtual serial port COM4 and to re-direct these messages to the line 25 of the screen, making them available for any other system.

The Italian experiment was "off-line". The evaluation package consisted of the CD ROM, the Italian version of the questionnaire in printed version and a diskette with the same questionnaire in ASCII file, to facilitate the blind user in filling it out. 11 copies of this package were sent by mail to users in contact with IROE as testers for the different products of research projects.

Subjects

The testers were chosen among those with a personal computer equipped with CD ROM drive and Sound Blaster SB16 card. Unfortunately, only seven of them were able to install successfully the program and listen to the hybrid book. Four users encountered problems with incompatibility of the program with their software and hardware architecture.

Therefore, only 7 subjects replied to the questionnaire. They are all totally blind; their age ranges from 24 to 50 years, with an average of 36.7. One is graduate teacher, 4 are telephone operators with a diploma of high or vocational school and 2 are university students. 5 of them know at least 1 foreign language; English: 4, French: 2, German:1. All regularly use the speech synthesizer and they are listeners of talking books; one declared to have heard from 4 to 10 books, 3 have heard more than 10 while 3 are regular listeners.

Results and comments

All participants stated that the concept of a structured digitized talking book was clear for them as well as the structure of the book. Also, the graph describing the structure of the book (tree) was clear to all subjects, although the basic entities of the digitized book were not comprehensible for one or two people. Most likely, they found some difficulties in understanding some concepts because the introduction was lacking in detail, or because they did not pay attention to some suggestions given in the introduction. In fact, some functions of the program require some reasoning in certain procedures.

One subject gave a perfect evaluation to the browser, while another 6 suggested performing some functions in different way. Five subjects replied that after some hours of training, each service of the browser could be used well by them, whereas two subjects were only able to use a few services. In conclusion, all stated that they would like to use the system.

This evaluation emphasized the paramount importance of the development of this new form of talking book. Subjects were impressed by the quality of the human speech associated with the facilities provided by the browser to navigate inside the text. After testing the different functions of the browser, a list of suggested improvements was provided by the testers. With regards to the digital audio format, it was advised that different formats other than MS ADPCM should be realized, in order to make the system more flexible and utilizable in different countries where varying standards of digital talking book will most likely be adopted.

Conclusion

The structured digital talking book can be considered the most innovative result of the DIGIBOOK project. Its potential utility appeared evident during the evaluation with the blind users. The interest of blind people's organizations, such as the Italian or Hungarian Blind Union, confirm this conclusion.

However, there are some open problems: first, use of the system by a blind person who is not a regular computer user needs a simplified system with CD player, which can be realized on a standard platform or a dedicated device; second, the handling of the material, text in HTML-like format and audio files, requires too much man-time. The study and development of automatic or semi-automatic procedures of segmentation of the audio file according to the structure of the book appears necessary in order to make this technique truly worthwhile in its application to future digital talking books.

Fortunately, research has not yet been completed, and is continuing in the direction of automating synchronization of the text and audio files, developing simplified multifunctional CD-player, and improving the browser functions. Knowledge of digital talking books requirements [2] has been summarized by the Helios project, and the answers and suggestions provided by Hungarian and Italian testers provide many ideas for improving the browser.

Acknowledgments

The authors wish to thank the Copernicus and Tide offices of the EU who made it possible to create new applications for visually impaired people in the DIGIBOOK project. We also extend our thanks to the Committee for Technological Development Hungary (OMFB) for their support, which allows us to continue our research, and for all the support from the Hungarian Academy of Sciences.

References

[1] Arató, A., Barkaszi, M., Buday, L., and Vaspöri, T., "Hybrid Books for the Blind - A new form of Talking Books", *5th International Conference, ICCHP'96, 1996, Linz, Austria, R. Oldenbourg Wien München, 1, pp. 237-*248.

[2] "User Requirements for the Next Generation of Talking Books" Draft document for consultation By the European Blind Union (EBU) Expert Group on Next Generation Talking Books (EBU-TBEG)

"AN ELECTRONIC NEWSPAPER FOR THE BLIND" WORK SUMMARY

Amo, F.; Barreiro, J.M.; Caraça-Valente J.P.; López, G.; Navajo

Garrido, R.J.; Montes, C.

One of the many problems facing the blind in the information field is access to the written pr PEIN (Electronic Newspaper for the Blind) is a project designed to overcome this problem, givi this community access to the press by means of a conversion, transmission and reading syste Using this system, the blind now have access to *Diario 16* (a Spanish national newspaper) and wo is continuing on other newspapers, including *El Mundo* and *El País*.

The work presented here basically explains the PEIN architecture, defining the three mai subsystems. First, a description is given of the *Press Format Converter* which convert the newspaper from its original format into the PEIN format for transfer along the telephone line, als explaining the reasons for this conversion. Second, the *Newspaper Transfer and Storage* subsy is described. Its main function is to act as a newsagent (newspaper server), where the newsp are stored in such a manner as they can be accessed by blind users and then read from their PC Finally, the *Electronic Newspaper Reader* is discussed, which provides blind users with an eas means of accessing the written press using the necessary aids and appliances using for th tiflotechnical elements, as voice synthesis.

POSTER OVERVIEW

The poster presented seeks to address the major points of the PEIN (Electronic Newspaper for the Blind) project and is divided into three clearly separate parts. It is headed by the name of the application and its logo. The underlying architecture of the application is shown in the centre, where the basic operation of the electronic newspaper for the blind (PEIN) is illustrated by means of a diagram, identifying each phase and briefly discussing the most important features of its architecture and operation: the Press Format Converter, the Electronic Newspaper Transfer and Storage subsystem and the Newspaper Reader, the features particular to each subsystem being briefly described at the side of the diagram. Finally, the bottom part of the poster is divided into two separate regions, one showing the Centre of Computing and Communications Technology Transfer (CETTICO) logo and references (address, Web page, e-mail, etc.) and the other listing project partners (ONCE, national and local newspapers, magazines, etc.).

Hearing & Speech

Adamlab Re-Engineered: Low-Cost Digital Voice Technology for the Speech Impaired

Gregory A. Turner

Abstract:

This paper describes the Adamlab project which has pioneered an entrepreneurial model in the governmental (public agency) sphere for the design and delivery of low-cost assistive technology products to handicapped individuals for whom the otherwise high cost of this technology is prohibitive. Particular focus is given to Adamlab's re-engineered product line and the transition from synthetic to digital speech technology.

1. Adamlab- a Unique Project for Special Education Technology

Adamlab (Adaptive Devices Applied Methods Laboratory) is an innovative project for the design and manufacture of low-cost electronic and microcomputer technology for special education students who are mentally and physically challenged. ADAMLAB's speciality is voice output systems, in particular Voice Output Communication Aids (VOCA).

Adamlab personnel include engineers, programmers, and other technical and data management staff. Responding to the perceived and voiced needs of special education teaching staff, Adamlab designs equipment from the micro-component level up and then contracts with local assembly firms for manufacturing runs. Revenue from catalog-based sales nationally and internationally balance all expenses and net income is used to provide free equipment and services to students locally. By this, custom engineering services, repair and help desk type services are available to the local commimity at no cost to the taxpaying public.

2. Regional Educational Service Agency of Wayne County

Adamlab's parent organization is the Regional Educational Service Agency of Wayne County (RESA). The RESA is the level of educational administration and support services at the county level. Wayne County (Michigan, USA), which includes metropolitan Detroit and its southern and western suburbs, is one of the nation's most populous counties, with a school-age population of over half million. As an example of services provided, RESA's Computer Center operation supports the thirty-four local school districts of Wayne County for their finance and student record needs- one of the largest such operations for education in the US.

The Computer Center's field engineering service provided the infrastructure for the RESA's special education technology activities in the late 1970's. At that time, RESA was collaborating with Michigan State University's Artificial Language Laboratory (under the direction of Dr. John B. Eulenberg, one of the co-founders of ISAAC) in the development of synthetic voice applications for students in RESA's programs for the Severely Mentally and Severely Multiply Impaired. In 1978, the engineering program was transplanted entirely to RESA, eventually maturing into the Adamlab project in 1986.

3. The "Wolf" - Pioneer Low Cost VOCA

From 1978 to 1984, several "generations" of talking devices were prototyped, resulting finally in the development of the "Wolf" VOCA. Hand-held and battery-powered, the Wolf has a membrane touchpanel, on which overlays are positioned bearing pictures and symbols meaningful to the student user. Pressing the pictures activates the corresponding synthetic voice message (the "vocabulary") —so, for example, a picture of a television could have the voice correlate: "I want to watch TV." The Wolf's electronics consist of the Adamlab-designed single board 8-bit microcomputer, with banked ROM and EEPROM memory accommodating data storage requirements for programmed vocabulary and for the installed software synthesizer. This software synthesizer, a customization of the commercial *Provoice* from First Byte Inc, Torrance, CA, replaced the integrated circuit SC-01 in 1993. It has a slight kinship to diphone systems in that its atomic units are digital samples. One of the attractions of the Provoice synthesizer system is its extensibility to other languages (i.e. besides English), although this feature was never implemented in the Wolf.

The principal ideas behind the Wolf device are low-cost, durability, and (relative) simplicity of use. As a case in point, the Wolf's original $250 cost (1984) was some ten times less than the then $2,500 average cost of comparably used equipment from private manufacturers – and indeed, only since around 1990 are there now available commercial VOCAs in the one thousand dollar and under price range. Low cost made the Wolf available and practicable for use with all the speech disability groups, including those in the severe categories for whom the purchase of thousand dollar-plus devices will always be controversial. As of this writing, well over 15,000 Wolf devices have been distributed and total distribution of all Adamlab products is approaching 25,000 units.

4. Digital Voice in the Nineties

The nineties brought the dramatic technology shift of economical digital voice to the VOCA industry. Digital voice integrated circuit technology had been available for as long as ten years, since the mid 1980's but the high cost of non-volatile, write-able digital memory had militated against widespread usage of the technology. The typical application, battery-backed RAM memory, was also highly susceptible to chronic data loss. By the mid-nineties, however, the increasing sophistication of EEPROM memory technology facilitated a veritable explosion of VOCA products. Many of these provided the basic functions that the Wolf had been used for, with the appealing simplicity of operation inherent in the record/playback idiom (vs. phoneme "spelling" for synthetic voice technology) and at costs driven lower and lower by fierce competition. Virtually overnight, the market for synthetic voice equipment shrank to a fraction of its previous size.

At Adamlab, we instituted a major design effort to address this situation. The key to these efforts has been, to date, availability of a highly integrated circuit family from Integrated Storage Devices Inc (ISD). ISD microchip devices incorporate both digital codec and EEPROM memory into a single chip – the company also pioneered usage of n-ary (also known as pseudo-analog-type) storage in the EEPROM cell, allowing megabit devices to have eight-times the nominal capacity for the storage of speech data. Since the introduction of the first twenty-second part, ISD has regularly increased the data storage capacity and currently one- and two-minute parts are available in commodity pricing. ISD speech technology is paired with the highly integrated single-chip microcomputer devices from Microchip Inc to create the standard platform for Adamlab's current set of digital products.

The "Hawk" comes in a two-level, two-minute version and a three-level, three-minute version; ultra-sensitive touchpanel provides a two-row by four-column matrix for programming the eight messages possible on each level. The device also offers eight input jacks for replacement of one, several or all of the eight touchpanel cells by discrete single switches.

The "Blackhawk" is a four level, four-minute machine. It sports a downsized form factor to provide increased portability and is aimed at ambulatory clientele and users with more specific fine motor control. The four-by-four touchpanel (sixteen messages per level, sixty-four total) has a fifth column dedicated to level shifting.

5. Superhawk VOCA

The "Superhawk" is Adamlab's most recent product offering. It combines the best program features of the old Wolf VOCA with the convenience and utility of digital voice technology. Over the course of its tenure, the operating program of the Wolf evolved into a sophisticated set of routines and programming features that facilitated the use of the device into diverse assistive, educational and recreational contexts. These features have been ported to the Superhawk.

Auditory scanning allows the device to be used by more severely-impaired persons who do not have sufficient fine motor to exercise direct selection of the touchpanel: here, a single switch is used to control a scan of messages that are indicated by voice – typically muted and relayed by a second speaker mounted by the user's head for private audition – activating the switch then relays the message to the louder, so-called "public" speaker. The Superhawk incorporates a labeling feature for this auditory scanning that allows abbreviation of the message for the private speaker iteration; and it incorporates a sophisticated branching or categorial programming structure so that the messages can be arrayed into a logical "tree" structure for increased scanning efficiency. This same program feature can be used alternately for a kind of message encoding as popularized by the icon-based "compaction" type.

The various "game commands" allow the Superhawk to be programmed and used for recreational purposes. SPIN mimes a game spinner (or the roll of dice) by randomizing a selection from a set of messages. LIST pulls selections one by one from the set of message choices. These routines allow handicapped children to engage in board game type activities in a meaningful and interactive way. Clinicians have also discovered the use of LIST command in early-literacy education where it can be used to allow the child to "read" the text of storybooks which have been entered as a message set.

Like the Hawk, the Superhawk sports input jacks to permit attachment of single switches for touchpanel replacement. Four additional output jacks are relay outputs and can be programmed to respond to touchpanel or single switch to activate toys and other such devices that have been modified to be operated with relay closures. This feature allows the Superhawk to figure into early childhood, pre-educational classroom and home activities and can be employed in cause-and-effect, discrimination, and sequential choice training.

Vocabulary capacity is six-minutes or twelve-minutes, organized into up to seventy-two levels. Levels have a maximum resolution of six-rows by twelve-columns (seventy-two messages maximum per level) but the touchpanel can be programmed into fourteen different matrices to provide larger selection areas.

Navigating and manipulating the vocabulary "space" of different levels and multiple messages can be more-or-less sophisticated depending on the cognitive abilities of the end-user. The so-called GOTO command can be used to designate one or more selection cells as level shift actions. The automatic, background-tasking message buffer can be used to string individual words and phrases together for whole-sentence speech.

Feedback for programming and recording is provided by a two-line LCD display and by a tricolor LED. A volume control rounds out the set of features.

Near future enhancements of the Superhawk platform include a compressed speech module option for 60 minutes speech capacity and a visual scanning option that adds surface moimted LED's to the 6x12 touchpanel matrix.

6. Adamlab as an Economic Model

Adamlab products are priced to recover direct costs (components and the cost of contracted manufacturing and miscellaneous other "raw" costs like postage), associated

Adamlab staff costs and pro-rated estimated costs to cover loss, breakage, repair under warranty, on-going R&D costs, and costs associated with with equipment (e.g. computers) re-capitalization.

In addition, a charge is added to provide funding for Wayne County-specific special education activities. So, for example, in 1992 a site-based project, the Assistive Technology Resource Center (ATRC), was launched with funding from these designated funds. During this startup period, ATRC's evaluation, assessment, and consulting activities were carried out by a three person team, consisting of speech clinician and occupational therapist and a computer systems consultant. Besides assisting local school districts in activities such as IEP, inclusion, vocational education, and transition, ATRC staff also engage in research-related activities. In 1993, a collaboration between ATRC, Adamlab and students and faculty at Wayne State University's Department of Electrical Engineering resulted in the first place design win at the RESNA student competition. The design, called the Sequential Task Trainer, uses sensors and lighted and voice-output feedback to prompt and correct low-cognitive students in manual assembly tasks, such as those typically found in sheltered workshop contexts. This field, technology products for vocational education, is one of the exciting new areas of application for assistive technology and rehabilitation.

Adamlab's positive revenue situation permitted the provision of no-charge repair services from 1986 to 1996 when a minimal $25 bench fee was implemented – this flat fee covers any and all labor and parts. In an industry where battery replacement is typically a multi-hundred dollar repair cost, this kind of repair fee policy operates significantly to the financial advantage of the consumers who typically have not factored in repair to the cost of assistive technology products. Adamlab's ability to provide repair at such low cost means that long term maintenance of product by consumers is economically neutral to them.

Adamlab is a "win-win" situation; appropriate technology is provided at very low costs to adequately meet the demands of the consuming public, at the same time yielding revenues sufficient

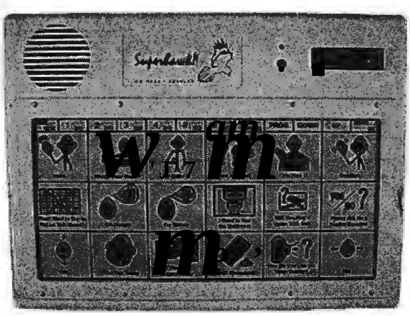
- for the long-term maintenance of the equipment,
- for providing custom engineering, free product and consulting services to local consumers,
- for supporting on-going research and development to practicably apply advances in technology to the needs of handicappers.

Hawk VOCA

- 7.5 seconds of digitized speech per selection area (two minutes total)
- Selection areas are 2" by 2"
- Two levels of eight selection areas, 16 messages total
- Sensitive touchpanel with no deadspots
- Built-in switch accessibility – eight jacks
- 2x4 keyguard included
- Volume control

Blackhawk VOCA

- 3.75 seconds of digitized speech per selection area (four minutes total)
- Four levels of 16 selection areas, 64 messages total
- Switch accessibility via peripheral adaptor
- Volume control
- Options: gridded keyguard, shoulder/wrist strap, easel/handle



Superhawk VOCA

- Six minutes total digitized speech
- 72 programmable cells w/ adjustability from 6x12 matrix to 1x1
- 72 pages limited only by number of messages
- Single switch auditory scanning with categorial branching commands and labeling
- Four direct select switch inputs
- Four programmable relay outputs
- Game commands
- LCD display for level location and programming functions
- Ultra-sensitive touchpanel
- 72 location keyguard included
- Volume Control

Adamlab VOCA Products: Hawk, Blackhawk, and Superhawk

INCLUDING INFORMATION AND COMMUNICATION TECHNOLOGIES IN THE TRAINING OF TEACHERS OF THE DEAF

Dr Magda Nikolaraizi, Mr Peter Fox

Abstract

This study discusses the current role of Information Technology (IT) in the education of deaf children by exploring the perceptions of lecturers offering training courses for teachers of the deaf and by eliciting the views of teachers working with hard of hearing and deaf children. Lecturers regarded I.T. as a positive educational intervention, however they did not distinguish the multifaceted dimensions of I.T in the training courses. Teachers of the deaf integrated IT into their teaching and students' learning, but they could not form an integrated picture of the potentials and implementations of I.T. in the education if the deaf. Under these circumstances, this study capitalises on the growing realisation of the need to clarify the role of IT and integrate it in the special training courses and in the education of hard of hearing and deaf children.

1. Introduction

The supportive role of Information Technology (IT) in the education of all children is now well established as a potential benefit within their development. The realisation of this potential will come to fruition through a proactive approach from teachers, who have a commitment to realise the benefits for the pupils in their care. As part of an on-going review of the incorporation of IT into teacher training courses the particular role of IT for deaf children was seen as a strong concern.

The supportive role of computers for pupils with special educational needs is well documented [4] and has also been acknowledged through the establishment of research centres, which were set to explore the role of IT in the education of children with special needs, including CEMEC (The Centre for Micro-Assisted Communication), the four SEMERCs (Special Education Micro-Electronics Resource Centres) and the ACE (Aids to Communication in Education). A more focused initiative in the area of deafness was observed in 1989 when The National Council for Educational Technology (NCET) created a specialist group. In parallel, The British Association of Teachers of the Deaf (BATOD) formed a 'Committee on Technology' resulting in a joint BATOD/NCET conference in 1991 which emphasised the important role of IT for deaf children [6].

Despite the growing appreciation of the role of I.T. in the education of the deaf, it is questionable whether this role is fully appreciated by the professionals who are mostly involved in the education of hard of hearing and deaf children, namely the teachers of the deaf. Furthermore, teachers' understanding and awareness regarding IT cannot be seen in isolation from their training and in particular the special training that they receive. Therefore, we believe it is essential to examine the involvement of IT in training courses for teachers of the deaf.

The agreed document of the training establishments in the UK, which provides a unified perspective on the nature and content of courses leading to the mandatoiy qualification of teacher of the deaf, presents Computer Assisted Learning (CAL) as the only contribution of IT to the special curricular needs of deaf children.

> "Does a one year full-time course or its equivalent really provide the depth of knowledge and understanding for teachers of the deaf to feel confident and competent in their multi-faceted roles? I believe a basic foundation can be provided, but at the same time think that serious discussions need to take place to review the time span." [14, p.V]

The above document specifies that during their training, teachers of the deaf can only expect to acquire a basic foundation on the multiple roles expected of them in their work with deaf children. As a consequence, the roles of these teachers are prioritised and IT seems to be discarded because other roles are seen as "more important". A strong initiative needs to be developed, though, ensuring that qualified teachers of the deaf will be well informed across a wide range of issues including the full awareness of the role of IT.

In the light of these beliefs teachers of the deaf and trainers from institutions were approached and their views on IT were elicited.

2 Methodology

To ascertain the feelings and understanding of teachers of the deaf on the role of IT, evidence was gathered by questionnaires from both lecturers and teachers of the deaf. Aiming at extracting understandings of the major issues surrounding the role of IT in the education of deaf children, the questionnaire was semi-structured encouraging free comments that could inform us on the beliefs of the respondents.

When constructing the instrument the term IT representing Information Technology together with a separate indication of Communications Technology was used. This was done with the concern that introducing another new acronym, ICT, would tend to confuse our respondents as the term had only just appeared in the UK.

After an initial precontact with the lecturers by telephone the questionnaire was sent to all the 7 training institutions in England and Scotland. Responses were received, even after subsequent telephone contact, from 5 out of the 7 institutions (response rates 71%) offering training courses for teachers of the deaf. To elicit the teachers views 35 schools, including units and special schools, were randomly selected from the National Deaf Children's Society (NDCS) directory and covered schools from England and Scotland. Twenty anonymous responses were received (response rate 54%) from qualified teachers of the deaf, eight from units or resource classes from special schools for the deaf.

All individual views were analysed according to the following way: Teachers' and lecturers' responses were segmented into relevant and meaningful units [13], which were subsequently grouped together according to the following categories: a) IT in teaching and learning, b)

Considering the special needs of deaf children and c) IT use in the language curriculum. Recognising the strength of the lecturers' and teachers' comments numerous participants' responses were quoted, a method that has been applied before [9] aiming in this way to bring the readers into direct contact with the research findings and present the data *"in such a manner that the informants speak for themselves"* [12, p.21]. The participants' comments were further enriched with the authors' interpretations, supported by their professional and personal experience and further analysed based one the review of the literature.

Individual comments were not personally attributed as the small number of institutions involved would allow the identification of our respondents. For the same ethical reasons personal backgrounds of the participants, although collected have not been used in any of our analyses.

3. Findings

3.1 General role of IT in teaching and learning

Lecturers comments illustrated the low attention that IT has received in the training courses. In two training courses IT was not addressed at all, while in the rest three training courses IT was addressed only in I.T. sessions which were offered mostly by invited lecturers. The majority of the lecturers acknowledged the positive role of such sessions in highlighting the benefits of I.T. in the development of literacy in deaf children. One lecturer, though, refused to see any benefit deriving from such sessions by saying that... *"IT is not very helpful because all students are at different levels"*.

Furthermore, these sessions presented only narrow-sided information and not well structured knowledge of the meaning and the potential of IT, failing to enable teachers to acquire a broad understanding of the role of IT in the education of the deaf. The limited consideration of IT was illustrated in the following two representative comments, where IT was associated only with literacy and speech development or only with sign language and Bilingualism. Lecturers did see a positive side of IT use but they appreciated only a partial contribution of IT to support their teaching and students' learning, which may also express their beliefs concerning the most appropriate mode of communication for hard of hearing and deaf children.

"We have sessions at the residential school for distance learning students and for full-time students on educational software to teach literacy, speech".

"Some sessions on the use of a) Sign Graphics b) Bilingual storytelling".

Furthermore, lecturers demonstrated through their comments that they did not actively integrate IT in their lectures as only 1 lecturer from one training course referred to the following limited use of IT in supporting his teaching and students' learning *"...preparation of materials, overhead handouts..".*

Numerous interpretations can be made of these results, including that teachers were negative towards computers, that they did not feel confident with their use or that they were not aware of their personal role. Microcomputers have been described as the most anxiety- and fear-producing innovation among teachers in the past two decades and teachers' lack of knowledge has been associated with unconfident attitudes and limited use of IT in the education of the deaf as has also been advocated by Mertens and Rabiu [8] The corollary then is that if lecturers do not feel and demonstrate confidence with IT, they will affect the attitudes of teachers of the deaf, which will mediate against the integration of IT into the education of deaf children.

Lecturers failed to report any type of special support that IT could offer teachers of the deaf in teaching and learning, making comments such as the following: *"IT is useful for all teachers"* A lecturer from one training course, though, referred to the following I.T. implementation for teachers of the deaf in bilingual settings *"IT is useful to combine BSL on video with equivalent English texts and present in this way bilingual stories".*

Teachers of the deaf in this study believed that IT played a supportive role in their teaching as well as in the learning of deaf children, but when asked to explain in what way they made use of IT to support teaching and learning their comments were perfunctory with two respondents failing to offer any comment at all. On several occasions (7 teachers) IT was considered as an assisting tool in teaching through the creation of worksheets ... " *making worksheets, mouse and control skills and storing/retrieving work"*, the role of which in developing spelling, comprehension, creative skills and providing deaf children with plenty of experience of meaningful language has been well established [2]. Only a secondary teacher offered a more detailed description concerning the role of IT in her teaching *"Revision materials created by me to support parts of the curriculum, My World software and multimedia and I also make videos to teach with".*

Most of the teachers' comments concerning the impact of IT on pupils' learning gave the notion of lack of serious thought such as the following representative following comment..."*IT is useful to all teachers and is part of the National Curriculum*". Such comments deny any understanding that the computer should not be considered simply as part of the curriculum, but more as a device for generating and modifying that curriculum offering new approaches for teaching and learning [11].

Teachers linked the positive role of IT in learning (10 teachers) to the use of particular software recording a list of computer software by generic title, including in descending rank order wordprocessing, electronic communication, graphics, multimedia, spreadsheets and speech software. However, the majority of respondents could not report how they made use of the software programs to support various parts of the curriculum, which poses the following question. How could teachers use effectively the software programs since they failed to see how these programs could be implemented into their educational practice? Although there are software programs which almost determine the sort of activities that accompany its use in the classroom, there are also other software types which are neutral, being highly malleable and responsive to various teaching approaches. Therefore, as others have established, teachers need to be informed and retain a great deal of control over how software programs and computers will be used, in order to adapt them and interpret them to fit their own philosophy of education and the best interests of the learners for whom they work [11].

3.2 Considering the special needs of deaf children

Teachers' responses indicated a real concern that the power of the computer to support pupils in what they are taught and in their learning has not been fully disseminated. When teachers were specifically asked in what way could IT support deaf children, 4 teachers did not differentiate the role of IT in the education of the deaf from its use in general education, responding in a vague manner: "*...to move forward in this society every teacher must be computer literate*" or giving a list of activities such as "*...record keeping, CAL programmes*" ...without any further comment concerning how these activities might correspond to the needs of deaf children.

Other teachers who gave positive responses and describing computers as "*very visual*" which reflects what has been described as being particularly useful for deaf children, mainly as it helps them to process and accept information in a mode that presents the minimal barriers to the child [5].

A small group of teachers indicated an understanding of the needs of deaf children and the way that IT could be implemented to respond to those needs saying that IT" *...allows for repetition in a new context and provides experiences which are otherwise not available*". This illuminative comment clearly reflects a well established knowledge, that hearing loss limits part of the experiences that children may acquire through the auditory channel [7] and emphasises the role of I.T. in enriching the experiences of deaf children.

3.3 IT use in the language curriculum

Although lecturers did not strongly comment on the positive role of IT in the training of teachers of the deaf, they did stress that IT facilitated access to the curriculum recognising its special role in language development, commenting as follows: "*IT offers opportunities to practice and improve reading and writing skills as a means of communication additional to Sign Language*".

Half of the teachers stressed IT's reinforcing role in the English curriculum. They argued that IT encourages listening skills, facilitates literacy, learning of grammar, spelling and writing. Only 2 teachers regarded IT as an assisting tool for the acquisition of signing skills and the positive role of IT in teaching signing and fingerspelling, which has been strongly promoted by Wilkins [16]. The association of IT with language development did not surprise us, since much has been written on the value of IT to support language learning including the use of word processors as a written language intervention and also by allowing children opportunities for interaction, initiating language or taking part in communication [15].

Two teachers, one of whom, was also a mother of a hard of hearing child indicated the positive support offered by electronic communication in language development. The significant role of such interaction towards the development of writing skills has been seen through the use of electronic communication, as for example in the USA, where students typed messages to fellow students and in this way improved their writing by interacting and communicating with each other [6]. Also, five teachers acknowledged the positive role of Internet style services recognising some of their dimensions "*access to information*" or "*increased circle of friends*"

It is rather disappointing that the role of interactive video was absent from the teachers' responses although it can combine the power of video materials with the interactivity afforded by a computer. These resources can be enormously helpful because they constitute purposeful activities that motivate and stimulate children and in this way enhance their language and communication skills [3].

4. Conclusion and Discussion

IT has a positive role in the education of deaf children and as has been explained teachers of the deaf nowadays do not receive adequate training concerning IT in their specialist training. Although it cannot be ignored that teachers may have the opportunity to receive later in-service training and increase their knowledge concerning IT, it is more likely that after finishing their special training they will be unaware of the full implication of this role.

This has been traced amongst the teachers that participated in this study, who considered IT as a positive intervention for teaching and pupils' learning and as being particularly useful for teachers of the deaf, but when asked to justify their responses they did not present a clear understanding of the role of IT in the education of the deaf. *Although they did* recognise the potential of IT regarding language work, they had limited views on the various ways that language could be promoted through IT. They acknowledged the positive role of word processing but the interactive features of IT towards language development through electronic communication or interactive video have not been substantially recorded.

Moreover, lecturers did not appear to use IT in their teaching, presenting in this way the role model that they did not appreciate the potential of IT in student's learning and thus excluded IT from training courses due to other priorities. IT must be included in the proactive development of the many roles on training courses and it is essential that consideration be given as to whether criteria for courses need to be changed to place emphasis on the importance of the role of IT in the education of deaf children. It has been well documented that a special needs educator, such as a teacher of the deaf, needs to be knowledgeable concerning the best ways to maximise the potentials and the abilities of their students to interact in a wide range of environments and since IT plays a basic role in enabling teachers to accomplish this task and a teacher of the deaf needs to be confident their IT competencies [1].

It is therefore paramount that all courses for teachers of the deaf should include the role of IT (or ICT) in the development of teaching and learning with special emphasis on language development. Training in the use of specific software we feel is insufficient, but an ongoing process of consideration and reflection of how and when IT can effectively support their pupils' learning must be investigated for these teachers. Lecturers need to adopt and develop the role of IT to ensure that the teachers receive the benefits of this understanding but also to encourage the use of IT in their own professional work, thus establishing a positive role model.

The demands that such an initiative would have on training courses, requires further investigation and the realities behind many of the comments of our respondents need closer scrutiny. It is immediately desirable to follow up on the lecturers and teachers to allow them greater time to explain their concerns and support them in identifying their needs towards the use of IT on their courses and in their classrooms. We feel that deaf children deserve this time so that their experiences and expectations can be enhanced by effective use of ICT in their education.

Notes and Acknowledgement

This study was based on the responses of the teachers of the deaf and the lecturers of the training courses for teachers of the deaf. Their support and interest in this research is acknowledged and we would like to thank them for their willingness to participate in this study.

References

[1] BEIGEL, A., Developing Computer Competencies among Special Needs Educators. Learning and Leading with Technology, 23, 69-71 (1996).

[2] GRAY, D., Computer Assisted Learning and Hearing-impaired Children: Part I; Does CAL Work? The Journal of the British Association of Teachers of the Deaf, 19, 38-46 (1995).

[3] HANSON, V. and PADDEN, C., Interactive Video for Bilingual ASL/English Instruction for Deaf Children. American Annals of the Deaf, 134, 209-213 (1989).

[4] HAWKRIDGE, D. and VINCENT, T., Learning Difficulties and Computers: Access to the Curriculum. Jessica Kingsley, London 1992.

[5] HUTSON, B., Adapting Computer Software and Hardware for Students with Special Needs, Advances in Spcial Education, 6, 171-188 (1987).

[6] LEGO, G., An investigation into the Use of Computers with Hearing-impaired Children with a Particular Focus oi] their Role in Aiding the Language Development of Secondary Aged Children, Unpublished Med Dissertation, Sch of Education, University of Birmingham (1995).

[7] MARLATT, A., ENFI-An Approach to Teaching Writing through Computers, American Annals of the Deaf, UlJ 240-244, (1996).

[8] MARSCHARK, M., Psychological Development of Deaf Children, Oxford University Press. (1993).

[9] MERTENS, D. and RABIU, J., Combining Cognitive Learning Theory and Computer Assisted Instruction for Deaf Learners, American Annals of the Deaf, 137, 399-403 (1992).

[10] NIKOLARAIZI, M., Exploring Parents' and Teachers' Perceptions of Deaf Children in Greece., Unpublisl PhD Thesis, School of Education, University of Birmingham (1997).

[11] SCRIMSHAW, P., Computers and the Teacher's Role, in; B.Someck and N.Davis (eds). Using Informa Technology Effectively in Teaching and Learning: Studies in Pre-service and In-service Teacher Educa Routledge, London. (1997).

[12] STRAUSS, A. and CORBIN, J., Basics of Qualitative Research: Grounded Theory Procedures and Techniqi Sage, Newbury Park, CA, 1990.

[13] TESCH, R., Qualitative Research: Analysis Types and Software Tools, Falmer Press, London 1990

[14] TRAINING ESTABLISHMENTS FOR TEACHERS OF THE DEAF IN ENGLAND, NORTHERN IRELAND SCOTLAND AND WALES, Training Teachers of the Deaf: Content of Courses, University of Manch Manchester, 1995.

[15] WARD, R. AND ROSTRON, A., Computer-assisted Learning for the Hearing-impaired: An Interactive Writt Language Environment, The Volta Review, 85, 346-352 (1983).

[16] WILKINS, M. Look what Microcomputers are up to Now? Perspectives for Teachers of the Hearing-impaired, 5-8, (1989).

A SUPPORT SYSTEM FOR EXPLAINING DIAGRAMS AT THE CLASS OF HEARING IMPAIRED STUDENTS USING A LCD PEN TABLET AND A VIDEO PROJECTOR

Hiroki Minagawa[1)]

1. Introduction

Diagrams are very important media that can express several information as spatial relations. It is, however, difficult to understand the concepts only to read it. If there includes something unknown concepts, it is necessary to be given some explanation.

When someone explains a diagram to the other one(s), the expositor may point on the focus of the diagram using a fingertip, a stick and so on, then adds the linguistic explanations as voices. The pointer sometimes moves on a point, sometimes on an area, and sometimes moves along with the story made by the expositor. The linguistic explanations are linked to the points, the areas, and the stories of the diagram.

It looks like multimedia consists of visual and auditory representation[1],[2],[3].

Diagrams are very important media for hearing impaired because of the excellent visual representation. It is, however, difficult for hearing impaired to be given the explanation via voice.

This study aims to support for hearing impaired to obtain the explanation of diagrams, especially in the class of the hearing impaired students.

2. Concepts

A diagram a lecturer made for the lecture drawn on a sheet of paper is explained by him/herself or an assistant. The diagram is captured through a video camera, then it is projected to a screen each students can see from anywhere. The lecturer or the assistant can insert letters by their voice, or hand writing letters, or typing keyboard. The letters are related to the appropriate location on the diagram using a pen tablet display or a mouse. The focus along with the lecturer's story can express as a point on an area using such pointing devices. The explanations can easily be erase from the diagram if it is necessary, so the focus of the lecturer's story can be seen clear. It is necessary to cope with some ad-lib explanation by the lecturer or some question by the students rapidly using hand writing explanation. The diagram with the explanation that the lecturer or the assistant has made can serve to the students as the handout through LAN.

1) Tsukuba College of Technology. 4-3-15. Amakubo, Tsukuba, Ibaraki 305-0005 Japan

3. Concluding Remarks

A support system for explaining diagrams for hearing impaired is suggested. It is necessary to realize this system along with this suggestion as soon as possible.

This study referred to the way to add explanation to the bitmap image of the diagram. By separating the input of diagram image, the representation will be more flexible. This study includes a clue to realize a new communication media for hearing impaired. That is, there will be a new communication media based on the pen tablet display, and the representation is closed for hearing impaired users. If this pen tablet based system would spread to each hearing impaired user, the diagrammatic interactive communication at a distance is realize through LAN.

This research is supported by the Ministry of Education, Science, Sports, and Culture (grant-in-aid for scientific research #09780164).

4. References

[1] H. Minagawa, N. Ohnishi, and N. Sugie, "A Vision Substitution System for Reading and Writing Diagrams Using Tactile and Auditory Senses", Proc. of 3rd Int. Conf. on Computers for Handicapped Persons, pp. 353-362, 1992.

[2] H. Minagawa, N. Ohnishi, and N. Sugie, "Tactile-Audio User Interface for Blind Persons", Proc. of 4th Int. Conf. on Computers for Handicapped Persons, pp. 569-574, 1994.

[3] H. Minagawa, N. Ohnishi, and N. Sugie, " Tactile-Audio Diagram for Blind Persons", IEEE Transactions on Rehabilitation Engineering, Vol. 4, pp.431-437, 1996.

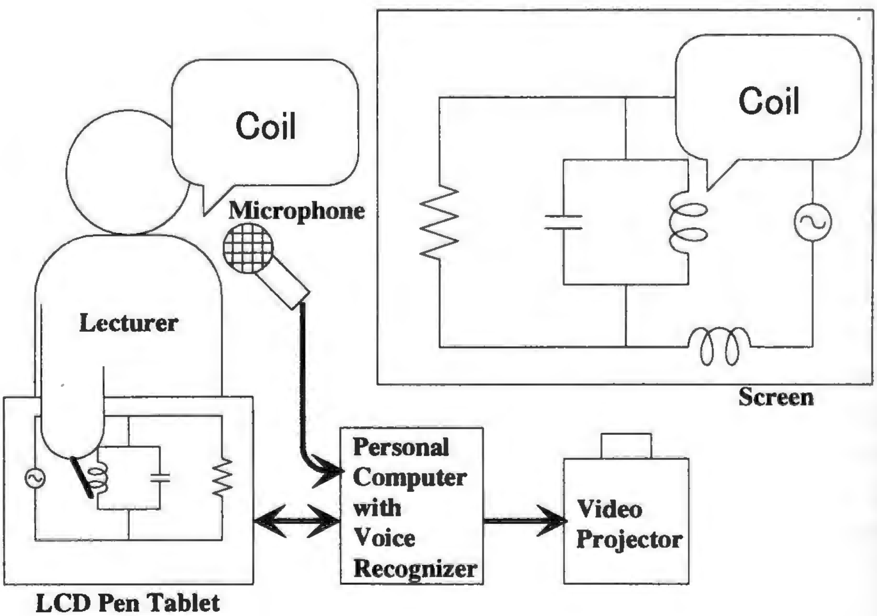


Figure 1. System Concepts

Session III

Non-visual Communication for Blind People

NAVIGATION AND INFORMATION SYSTEM FOR VISUALLY IMPAIRED PEOPLE

I. Kopeček and P. Smrž

Faculty of Informatics
Masaryk University
Botanická 68a, 602 00 Brno
Czech Republic
tf.:+420-5-41512364
fax:+420-5-41212568
e-mail: {kopecek, smrz}@fi.muni.cz
http://www.fi.muni.cz/~kopecek/
http://www.fi.muni.cz/~smrz/

Abstract

Orientation is one of the most important problems of visually impaired people. The aim of this paper is to suggest a contribution to the solution of this problem using computer technology. The basic idea is the detection of motion and orientation using sensors and consequent position identification. The detected trajectory is compared with a map and is corrected by means of the algorithm described in the paper. Some problems concerning sensor detection of human motion are also discussed. Based on the determined position other relevant information is provided to the user of system (information describing the neighbourhood of the actual position, optimal way to the chosen destination, possible warnings).

1. Introduction

Contemporary information technology offers various possibilities for solving many problems of handicapped people. Serious impairment of sight is a considerable limitation for such people. The main problem is a substantial information deficit and its feedback - communication barrier. Under normal circumstances about 80% of the processed information is mediated by the dominant sense - sight, which is only inadequately substituted by touch, hearing and other senses.

One of the most important problems of visually impaired people is orientation. This problem is usually solved in such a way that sightless people learn only several basic routes. Their orientation outside these routes is difficult without additional information and can be, to a certain extent, even dangerous.

Despite the growing attention paid to the problem of the orientation of blind people during the last years (see [1], [2], [8]), the problem is not sufficiently solved yet.

The main goal of this paper is a presentation of some principles of navigation and information system for sight handicapped people. The purpose of the system is not to solve all the problems involved in orientation and route control of blind people but to serve as an additional tool which can help sightless people respecting their psychology and social feeling.

Our idea is based on the following forethought:

- State authorities own detailed town maps in a digital form of GIS (Geographical Information System) and they are likely to agree with their use for orientation of sightless.
- Multimedia sensors connected to computer system allow the orientation to be determined. This feature can be used for trajectory estimation.
- It is possible to determine the position comparing the digital map in the computer memory and the calculation of trajectory of sightless using self-correcting algorithm.
- It is also possible to find a certainty factor for the position determined and to tell to which extent this position is reliable.
- In order to help the self-correcting algorithm to avoid situations when position determination is unreliable, estimating of the position by means of radio beacons on the frequency assigned for this purpose (88,790 MHz in the Czech Republic) and by the global estimation of position by GPS (Global Positioning Satellite System) can be used.
- Graph algorithms allow the optimal path to be found from a given place to the chosen one (the shortest way need not be optimal from the viewpoint of safety).
- Using speech synthesis it is possible to inform the user about possible roadblocks and dangerous places and to pass him (requested) information about his environment.

The system can work in two different modes:

1. Navigation mode is based on the map. The system navigates and informs the user about the actual environment.

2. Learning mode is used when no map data are available. The system memorises the trajectory that can be used in future processing as a part of memorised map.

For visually impaired people it is advantageous to choose a way that passes as few as possible crossroads and other dangerous places. For this purpose the length of each path section is weighted by safety factor. Then the common shortest path algorithm can be applied to find the optimal path.

1. Position, Itinerary and Map – Formalisation

In this section we formalise the notions of position, itinerary and map. This is suitable for use of formal apparatus and for exact formulation of some intuitive terms. We will consider an itinerary as an indexed system of positions after each step. Each position in an itinerary can be described by system of attributes. For the real implementation of system such attributes as determined physical position, global or local changes of direction, sound environment, terrain characteristics, etc., can be used. From this point of view, the set of all positions forming an itinerary can be considered as a special case of Pawlak information system.

Let us briefly recall that Pawlak information system (see, e.g. [5]) is an ordered quadruple S = (U, A, V, f), where U is a set of objects, A is a set of attributes, V is a set of attribute values and f is a mapping of the set UxA into V. The sets U, A, V are supposed to be finite.

Definition 1. A Pawlak information system **P** = (P, A, V, f) with indexed set P is said to be *itinerary*. More formally, we suppose that a bijection I from P onto the interval (1, ..., n) (where n is cardinality of P) is defined. The elements of P will be called *positions*. The position P indexed by i (i.e. the position for which I(P) = i) will be denoted by P(i).

For $1 \leq i \leq n$ we denote **P**(i) the itinerary derived from an itinerary **P** by reduction of set P to the set {P(1),... P(i)}. (This means, that **P**(i) forms a "subpath" of **P** from starting position to the position indexed i.)

The theory of Pawlak information system can be used e.g. for reducing of possible redundant attributes and for application of some algebraic methods ([5], [6], [7]).

Definition 2. A *map* is a pair M = (Z, d), where Z is a set of itineraries with common set of attributes and d is a metric on the set of all possible positions (for any itinerary).

For real implementation, d can be used in the form of weighted sum

$$d(P_i, P_j) = \sum_{a \in A} c_a d_a(f(P_i, a), f(P_j, a)),$$

where d_a is the metric associated to the attribute a and indices i, j are used to distinguish positions P_i and P_j. Suppose for instance that an attribute is of the type of local terrain characterisation; then the metric associated to this attribute should express the measure of similarity between instances of local terrain characterisation (value of the attribute of such type can be more complex than a real number, e.g., it can be a vector or matrix).

The weights α_i can be determined by heuristics or they can be obtained by experimenting and consequent evaluation of the optimal weights (for experiments done) by optimisation methods.

Determination of the values of position attributes for a map can be performed mostly automatically by computing the values of attributes from GIS maps (e.g., co-ordinates of position, global orientation, terrain characterisation, etc.,) and partly by experimenting with the system in the learning mode. If some data are incomplete the number of attributes can be reduced for the time before completion.

3. Algorithm for Computed Position Correcting

User trajectory is computed from a starting point at each step by determining the physical position of the feet by sensors. Comparing the attributes of actual position with the map can essentially reduce error that appears in this process. Unfortunately, the actual position of the user may not correspond to the path position (in the map) with the same index because step length can vary and because of physical position error. We use a modification of DTW (Dynamic Time Warping) algorithm (see, e.g. [4]) to overcome these difficulties. Let P_1, P_2 be two itineraries belonging to map M. We want to find a correspondence $f \subseteq \{P_1(1), \dots, P_1(n)\} \times \{P_2(1), \dots, P_2(m)\}$ which minimises the expression

$$\sum_{(P, P') \in f} d(P, P') \qquad (2)$$

by the following assumptions:

1. The correspondence f includes the pairs $(P_1(1), P_2(1))$ and $(P_1(n), P_2(m))$
2. If $(P_1(i), P_2(j)) \in f$ then exactly one of the following cases holds:

(a) $(P_1(i+1), P_2(j)) \in f$
(b) $(P_1(i), P_2(j+1)) \in f$
(c) $(P_1(i+1), P_2(j+1)) \in f$

Minimal value of the expression (2) under these assumptions will be denoted $D(P_1, P_2)$. (This value can be obtained by non-linear programming methods.)

It is easy to see, that

$S1(x,y)ef\} = \{P_1(1),\ldots,P_1(n)\}$

and

$|y|\{(x,y)ef\} = \{P_2(1),\ldots,P_2(m)\}$

Now, let P_1, P_2 be itineraries belonging to a map M = (Z, d) and suppose P_2 is the chosen "sample" itinerary whose physical positions were determined from a (physical) map. The user chooses this itinerary as the path he wants to follow. Suppose P_1 is the path that was determined by the system from the starting position to the position indexed by i (i.e. position after i-th step). For the itinerary P_1 the attribute of computed physical position is not precise and should be corrected. The correction is now obtained by the following steps:

1. Determine the index j, $(1 \leq j \leq m)$ which minimises $D(P_1(i), P_2(j))$.
2. Substitute the value of the computed physical position attribute of the position $P_1(i)$ by the analogous attribute value of the position $P_2(j)$.

This application of the DTW algorithm is motivated by the analogy with application of DTW algorithm for comparing spoken words (represented by sequences of acoustic vectors) and shows good practical results.

3 Step Detection by Sensors

In this section we would like to illustrate some problems arising in sensor detection of human motion. We are searching for methods that are suitable for determination of a particular stage of each step, e.g. the moment of treading (i.e. when the foot is laid on the ground, risen, etc.). These methods could be based on the following principles:

- Detection of the deflexion from horizontal plane
- Detection of the side deviation from the vertical plane perpendicular to the walking direction
- Detection of sensor vibration at treading

The detection of the deflexion from horizontal plane can be used for step detection if the sensor (Virtual i-O 3D Sensor) is attached to tarsus. In Fig. 1 the change of the deflexion is shown for this case. When the foot is risen at walk the deflexion increases till the maximum value is achieved, it is changed in a particular way during the course of the step, and, finally, depending on treading style it returns from the maximum positive or negative value to the starting position.

It may be concluded from our experiments that the use of this method is the most advantageous for detecting the time point of rising the foot from the ground. We are thus searching for places where the deflexion from the horizontal plane changes from zero value to maximal negative value. The most serious problems of this method are connected with shambling which is characteristic for old people. Also "tiptoeing" may lead to distorted results.

Fig. 1. Change of the deflexion from horizontal plane

Side deviation from vertical plane perpendicular to the walking direction can be used for step detection in case when sensors are attached to the thigh or shin. In Fig. 2 the changes of deviation from the vertical plane perpendicular to the walking direction are given. In most cases the maximum value is achieved at the moment of treading on the ground. This is the reason why we seek the maximum value in the graphic representation of step course. The method does not give good results for "marching".

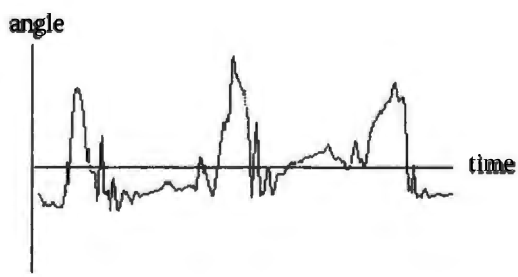


Fig. 2. Change of the deflexion from vertical plane

When both methods mentioned above were tested problems appeared which were caused by the fact that sensors could not be fixed on the leg tightly enough to prevent minor vibrations. In Fig. 3 the influence of these minor vibration on the change of the angle between monitored vector and reference vector (1,1,1) can be observed. Experiments showed that system behaviour in such case is rather regular.

This led us to attempts to use sensor vibrations at treading for step detection. In the graphic representation of angle changes we seek points near which the value is changing rapidly in all co-ordinates. The method works well in the case of quick paces, however, it gives wrong results if sensor vibrations are caused by other factors than treading.

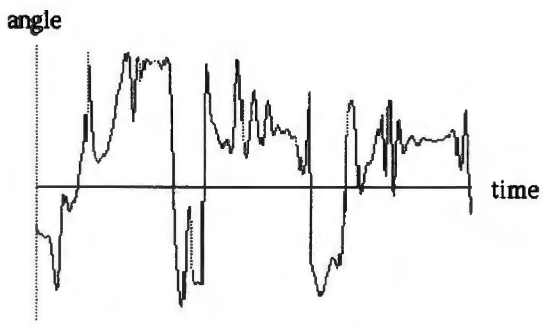


Fig. 3. Sensor vibration at strong treading

From the above discussion it can be concluded that no step detection method gives satisfactory results in all cases. For this reason we have chosen a combination of the methods, which is able to compensate a failure of one method by the reliable function of other methods. If one method detects a step a "voting" of all methods proceeds in which the vote of each method has a particular weight. These weights are generally dependent on the ground, walk characteristics, etc. At present we are trying to estimate weights in particular cases on the basis of experiments. We are also testing the utilisation of the learning mechanisms, especially neural networks ([9]), for automatic determination of weights from the experimental data.

4. Conclusions

The presented framework for trajectory determination and correcting combines the advantages of system that can learn trajectory with determining the position using GIS and GPS and correcting by other attributes detected or measured by sensors. Our present research concerns, besides the problem mentioned above, the topics of communication which should be carry out mostly by speech synthesis and recognition ([3]).

The further development of the presented framework will be directed toward experiments with precision and reliability of the sensors and building software components allowing to transform GIS data into a form suitable for system evaluation and testing the prototype of the system.

References

[1] T. AOKI, Y. YONEZAWA, K. ITOH, M. HASHIMOTO. Intelligent guide system for parametric sound beam for the visually impaired. In J. Klaus et al, editor, Proceedings of ICCHP '96, 637-644. Oldenbourg, Wien, 1996.

[2] D. D. CLARK-CARTER, A. D. HEYES, C. I. HOWARTH. The efficiency and walking speed of visually impaired people. Ergonomics, 29(6), 779-789, 1986.

[3] I. KOPEČEK. Speech synthesis of Czech language in time domain and applications for visually impaired. In Matousek et al, editor, Proceedings of the 2nd SQEL Workshop, pages 141-144, Plzen, 1997.

[4] C. MYERS, L. R. RABINER, A. E. ROSENBERG. Performance tradeoffs in dynamic time warping algorithms for isolated word recognition. IEEE Transactions on ASSP, 28(6), 623-635, 1980.

[5] M. NOVOTNY. On sequents defined by means of information systems. Fundamenta Informaticae, 4, 1041-1048, 1981.

[6] M. NOVOTNY, Z. PAWLAK. On a representation of rough sets by means of information systems. Fundamenta Informaticae, 6, 289-296, 1983.

[7] M. NOVOTNY, Z. PAWLAK. Algebraic theory of independence in information systems. Fundamenta Informaticae, 14, 454-476, 1991.

[8] H. PETRIE, V. JOHNSON. Evaluation methodologies for navigational aids for blind people. In J. Klaus et al, editor, Proceedings of ICCHP'96, 653-660. Oldenbourg, Wien, 1996.

[9] P. SMRZ. Hybrid connectionist-symbolic models - integrating neural networks and rule-based approaches, 1997 (manuscript).

Ultrasound Safe Way Detector with a Speech Output

RYSZARD KOWALIK, IRENA POSTAWKA

-Department of Optoelectronics, Technical University of Gdańsk,

80-952, Narutowicza 11/12, Gdańsk, Poland

e-mail: kowalik@sunrise.pg.gda.pl-

Abstract:

The paper presents considerarations of limitations in designing of mobile devices for the blind users. The principles of a model of ultrasound safe way detector worked out at the Technical University of Gdansk is presented. The method of determinig position of obstacles is discussed.

1. Introduction

The blind who move in an open space on their own are exposed to dangerous collisions with obstacles that are difficult to detect without the help of eyesight. These obstacles include poles, road signs, low hanging branches and cars parked on the streets. Many devices have been constructed to detect obstacles, however, they have not won the approval of the blind. What they actually did was making movement more difficult instead of helping to avoid collisions with obstacles in the way. An in-depth analysis of the features of these devices showed that they responded to the presence of obstacles without determining their position which made an adequate maneuver impossible in trying to avoid a collision without having to perform additional actions such as reaching out to actually feel the obstacle with the cane or hand. What is more these devices required a continuous tracking of the signals emitted. These signals engage the attention of the blind person and hamper the reception of sounds coming from the surrounding space. The end effect was that instead of assisting free movement in an open space, these devices made it even more difficult.

2. Needs and limitations of the blind users

Analysis of the possibilities, needs and limitations of the blind showed that the main barri encountered by a designer of a device meant to help a blind person with making their way across an open space is the slow speed of information flow through the other active senses. [1]. Therefore a device like that has to be complex enough to be able to take over the analysis of received signals in terms of determining the position of obstacles appearing on a blind person's way. Another task for the device is to figure out whether, and how, these obstacles could be avoided. The device shou also generate and convey to the blind user short verbal commands which explicitly describe the type of recommended action, for example, turn left, stop, take a strong turn right, etc. The transmis speed, other than through sight, of information to human consciousness is low. This enforces a reduction of the penetration area to the necessary minimum, i.e. to an area a little bit bigger than a corridor that is 1 m wide, 2 m high and 2 to 3 m long in front of the device which represents a sp that a blind person will walk across within the next several seconds.

To penetrate space the obstacle detecting devices use either visible or infra electromagnetic radiation or ultrasounds at the frequency of several dozen kHz. The relatively low frequency allows a digital processing of signals in real time and resulting from that, the possibil of eliminating undesirable phenomena such as fluctuation of echoes and accidental overlay of waves reflected from various objects which in end effect wrongly identify non-existing objects.

To convey information produced by technical equipment to the mind of a blind person one can only use the senses of hearing and touch. This is so because the other senses are underdeveloped. Touch with its wide reception field and many features that make it similar to eyesight, is potentially the best channel for transmitting two-dimensional information that the devices deal with. Unfortunately, so far we have not been successful in designing user frie cheap arrays of stimulators that would activate the receptors of this sense. Therefore hearing is the only sense capable of supporting the transmission of complex information with a practical mean for the user. At the same time one needs to be aware of the fact that hearing performs a major role in spatial orientation. Sounds heard by a blind person on the street help him to simultaneously obse many events happening around him. Correct recognition of the traffic, distance to vehicles, speed and direction of movement, sounds of vehicles braking before pedestrian walks or at public transportation stops and the reception of echoes reflected from buildings has such an elementa significance for a free and safe walking across the town that acoustic signals generated by the devi should not handicap the perception of the natural sounds of the environment. For this reason it was decided that the designed device should use short verbal messages informing about the next move to be made to avoid collision with an obstacle. Another problem to be solved and just as important is to transmit to the user's ears sounds of speech emitted by the device so that:

- the blind person could hear them properly,
- others would not hear them,
- they would not hamper the inflow of acoustic soimds from the surrounding space.

These three conditions can be met adequately and simultaneously if we use headphones with an open input which are used in some devices for people with impaired hearing.

The model of ultrasound safe way detector which answers the consideration mentioned above has been designed at the Technical University of Gdansk [1].

3. The method of determining position

To penetrate the space in front of the device ultrasonic pulses were used at the frequency of 18 kHz and duration of 1.1 msec. Using the beamforming method the signals received are transformed into 7 envelope signals corresponding to seven space observation cones. [3]. The cones whose vertexes are at the point of the device and have angles of 9 degrees can be placed horizontally (next to each other) or vertically (one on the other). The applied beamforming parameters show that the consecutive points of the envelope of received signals correspond to the points in space found on seven radii of the circular sector. The distance between the consecutive points on a given radius is 37 mm. At the same time it is the resolution of distance in this method. The area of the space undergoing the analysis was limited to a rectangular prism, 1m wide, 2m long and 2 m high placed 0.5 m (dead zone) before the device. (A bird's eye view of the situation is shown in Fig. 1).

[1] This work was supported by the grant no. 4 P05D 0303 of the Polish State Committee for Scientific Research.

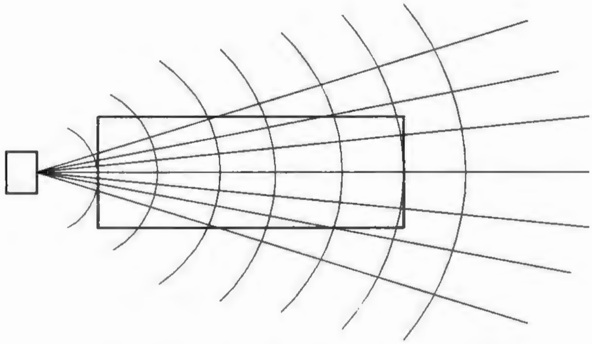


Fig. 1. The penetrating beams and search area.

In the course of the experiments it was found that apart from obstacles actually found in the device's field of view, it also detects non-existing objects which come up and disappear in the consecutive measurements. This undesirable phenomenon is the result of echo fluctuation and the interference of waves reflected from objects found beyond the area of penetration. This can be eliminated by the application of two types of echo amplitude summation – spatial (reduces the accuracy of object positioning) and so called progressive which sums up the values of envelopes from n recent measurements. By summing up the values of echoes from n measurements for a given envelope (where n= 2 to 10), an n times bigger area is treated as one point which reduces n-times the resolution. Since the goal of these efforts is to locate obstacles, i.e. searching in the received signals for envelope amplitudes with values higher than the experimentally defined threshold value, all numerical operations use unmeasured values of amplitudes, however, with values at 0 or 1 respectively for the value of the envelope that is smaller or bigger than the threshold value. Before the process of summation the modified binary values from n measurements are memorized in consecutive temporary maps. Another map stores values that have been summed up spatially and progressively. This is done in the following way. In consecutive measurement cycles (after sending a consecutive sounding signal) binary equivalents of amplitudes are added to another temporary map with full resolution and from n maps like this the summed up values are input into a map with reduced resolution. These operations are repeated in a cycle. To recap, in this method we use n+1 maps. In one of them a picture is created showing the location of actually existing objects while the other maps are used for eliminating accidental echoes. Naturally, this double system of summing of normalized signals does not reduce to zero the contents of those areas of the map which display false echoes exclusively. By selecting the right threshold value, though, we can obtain a stable picture of the area under penetration.

4 The block diagram and the principle of operation of the ultrasonic loud-speaking safe way detector

The block diagram of the ultrasonic loud-speaking safe way detector is presented in Fig. 2.

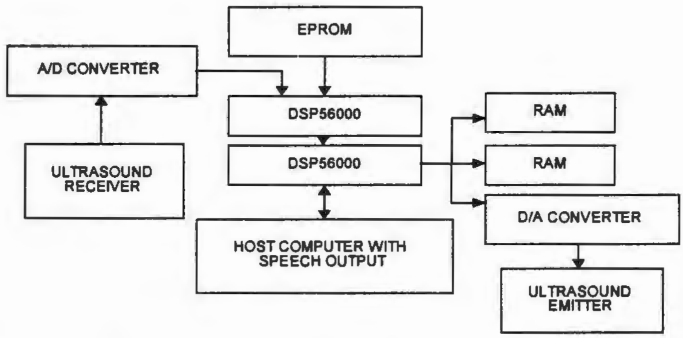


Fig. 2. The block diagram of the ultrasound detector.

The central unit is made of two Motorola DSP56000 signal transducers linked with each other through SCI serial ports operating in the asynchronous mode. They are also linked through HI ports with a host computer equipped with a speech synthesizer. The first DSP controls the analog-digital transducer A/D which processes signals from an ultrasonic waves receiver. The other DSP processor is equipped with two external memories of the RAM type and controls the digital-analog D/A whose output is connected with a high-frequency loudspeaker through an amplifier. In the SSI port of the processor a sampling frequency of the received signals is generated. In the address spaces of A ports of both DSPs permanent EPROM type memories are placed with a software that controls the operations of signal processors. [2].

The receiving part of the detector consists of an array of 8 x 8 electret microphones switched in cycles into 8 lines and 8 columns. The block diagram of the ultrasound emitter and receiver circuits is showed in Fig. 3.

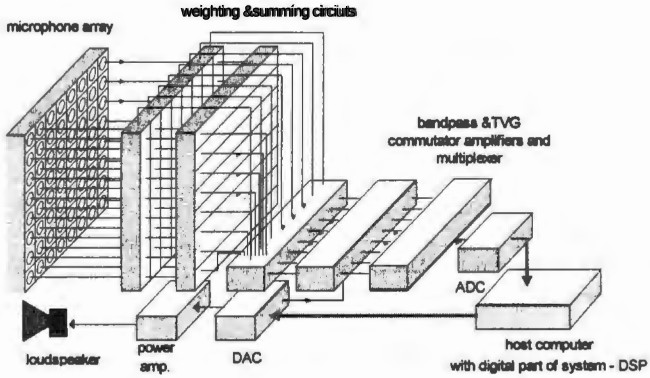


Fig. 3. Block diagram of ultrasound part of the system.

The signals from outputs of lines or columns through the weighting and summing circuits are passed on to the A/D transducer. When power is on, programs from the EPROM memory are automatically loaded into DSP and temporary tables from the host computer are loaded to the memories of both DSPs. In this way the system is ready to work. At a signal from the host comp the other DSP generates a sounding ultrasonic pulse whose duration is 1 ms which through a hig frequency loudspeaker is emitted into space in a solid angle of about 60 degrees. The wave reflected from the obstacles are received by the array of microphones and depending on the setti the switching array they are summed up with respective weights in columns or lines. Electric sig obtained in this way are passed on one by one to the A/C transducer input which conducts tw measurements of each of the signals every quarter of a wave period and transmits the meas values to the first DSP. This processor following the principle of beamforming transforms them in envelopes of seven signals corresponding to seven cones of space observation in front of the devic (Fig. 4.) Through the SCI serial port the values of the envelope are transmitted to the second DS where they are normalized and summed up following the above algorithm. The map created afte several echo signal processing procedures is analyzed for obstacle positioning.

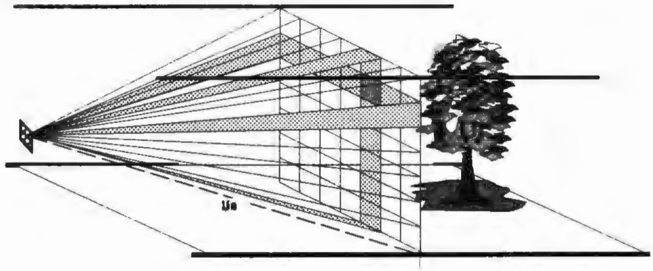


Fig. 4. The „view" of the tree in the horizontal and vertical beams.

If obstacles are detected in the received signals, the coordinates are calculated and sent to the host computer where the decision is made – what maneuver the blind user of the device is supposed to make to avoid the obstacle. An appropriate message is given through the speech synthesizer.

4. Conclusions

The presented model of the ultrasound detector of the safe way is a next attempt to solve a problem of a mobile device for the blind. The main goal of this device is to help the blind person to avoid the obstacles instead to signal their presence. Therefore a beamforming and signal processing method was used to process the ultrasound echoes and determind the position of obstacles. The speech output is used to give out a short clear commands to the blind user.

5. References

1. Kowalik R., Postawka I.: *Some aspects of reduction of information in sodars for the blind*, Proc. of the XIth Symposium on Hydroacoustics, pp. 211- 214, Gdynia (1994)
2. Salamon R., Kowalik R., Postawka I.: *Detection and signal processing in the acoustic echolocation system for the blind*, Archives of Acoustics, 22, 3, 1997, pp.277-295.
3. Salamon R., Marszal J., Raganowicz A.: *Multibeam echolocation system for the blind*, Proc. of the XIth Symposium on Hydroacoustics, pp. 267- 274, Gdynia (1994)

Non-Visual Access to Non-Textual Information through DotsPlus and Accessible VRML

John A. Gardner and Vladimir Bulatov

Science Access Project, Department of Physics, Oregon State University
Corvallis, OR 97331, USA

Abstract

Non-visual access to non-textual information presents many challenges. One challenge is that there are many examples of information in which the content and visual layout are almost inextricably tied together. When a clear separation of content and presentation is possible, a second challenge arises - there are few electronic formats that actually make this separation. A final challenge is that of finding efficient, intuitive, inexpensive methods of displaying nontextual information nonvisually. Two research programs at Oregon State University that address some of these challenges are described in this paper. One is the development of the DotsPlus tactile font set and TIGER (TactIle Graphics EmbosseR).
The second is use of VRML to produce audio/tactually accessible 2D figures.

Keywords:

DotsPlus, tactile graphics embosser, VRML nontextual information, print disabilities

Introduction

Speech or braille screen readers can provide excellent access to standard computerized text (although access can be severely hindered by complex electronic formats such as the PDF information provided for potential presenters at this conference). Very little other computerized information can be accessed easily (Barry 1994).

This paper discusses briefly two nontextual information access research projects underway at Oregon State University. Each is an attempt to find a general solution by defining or extending current paradigms for display or data storage.

DotsPlus

DotsPlus is a paradigm for representing text tactually in one of several fonts. Each character is unique and recognizable out of context. These characteristics permit DotsPlus to be used in graphics applications where text may appear in unpredictable locations. Examples include mathematical equations, and labels or other text notations on diagrams, maps, charts, etc.

The original DotsPlus font represented letters in 8-dot braille (Gardner 1993) Subsequently a number of improvements were made, in particular to the original punctuation marks, and a 6-dot DotsPlus font was developed (Gardner 1998). Lower case letters are standard braille in both font sets. A number of other symbols are also represented in braille but most of the thousands of symbols appearing in specialized literature are represented by tactile images of the ink print symbol.

In 8-dot braille and DotsPlus, capital letters are indicated by a dot in the dot-7 position (left bottom dot in a cell having four rows and two columns of dots). In 6-dot braille and DotsPlus, capital letters are indicated as a double width 6-dot cell. The right side of that cell is lower case braille, and the left side is the capital letter indicator. In English braille and DotsPlus, the capital indicator is a dot in the lower right position.

DotsPlus numbers are represented by European Computer Braille, and punctuation marks are represented by graphic symbols that feel much like braille punctuation marks but are distinguishable from dropped letters if examined carefully.

We believe that most literary braille readers who want to learn DotsPlus can do so rather easily. Apart from the numbers, standard text reads very much like uncontracted braille. A DotsPlus reader of more advanced literature must learn many new symbols, but this is identical to the learning process of sighted readers. A blind or sighted child must both learn the shape of a plus and equals sign before learning to do arithmetic for example.

There are only minor differences between DotsPlus fonts among languages sharing the roman alphabet. In 8-dot DotsPlus, the period (full stop) symbol shape may be altered for languages (e.g. German, Swedish) having a different braille period from the English/French symbol. For 6-dot DotsPlus the "prefix" symbol in the double cell braille symbols can be changed to reflect different capital letter indicators used in languages other than English. Such differences are minor enough that readers should have little difficulty reading literature printed in other languages. This is unfortunately not true for most braille literature.

Widespread testing and use of DotsPlus has been severely hindered by the difficulty and expense of printing DotsPlus. A wax-jet printer used for original DotsPlus research is no longer commercially available. Swell paper can be used to make DotsPlus materials, but swell paper is expensive and requires considerable expertise to make copy in which braille dots are easy to read.

A new technology invented at Oregon State University now allows embossing with resolution good enough to make DotsPlus materials. The TIGER printer, based on this new technology, is expected to become commercially available at a cost of approximately $6000 by the time of this conference. The TIGER includes a Windows 95 printer driver that permits direct printing from most Windows 95 applications. Users need to use a screen font with the correct size and should avoid complex multicolored or gray scale drawings. Otherwise virtually anything that can be printed on a standard printer can be printed on TIGER.

Accessible Graphics using VRML

Virtual Reality Modeling Language (VRML) is becoming popular in World Wide Web applications (Roehl 1996, Carey 1997). VRML allows one to create time-dependent three-dimensional models that can be displayed interactively. VRML "figures" are electronic files organized into a well-structured tree and are displayed by viewers that provide a two dimensional projection of the model (VRML 1997). Users may interactively modify the view by turning or moving through the model.

We have taken advantage of the power and flexibility of VRML to construct and display simple two dimensional figures such as those appearing in scientific literature at all levels. (Bulatov, 1998) We construct a VRML model using any convenient three dimensional objects whose projection is the 2D picture we desire. This may be done with standard authoring tools and eventually with a special 2D authoring tool we intend to write. This model is then modified with a special editing software application we have designed. With this editing software one may produce a second VRML file in which each object can be provided with a label that contains information that is, in principle, arbitrarily rich. Presently we permit only plain text.

Standard VRML browsers display the second file identically to the first. However if the model is well structured, and the labels are sufficiently informative, the second VRML figure is completely accessible to people with print disabilities through one of a number of specialized "viewers".

The special viewers are programs that supplement a standard VRML viewer by interpreting the special labels. The simplest special viewer, and the only one that is near completion at present, uses a common technique that permits a blind user to "read" a complex tactile figure with the help of a computer. A tactile copy of the VRML picture must be made and placed on an external digitizing pad. A blind user may then explore the tactile figure and request information about objects. This request, made for example by pressing on an object and activating the digitizing pad, causes the label to be displayed on the computer screen and, if desired, browsed in audio through use of an internal speech engine. In principle the label can be read with a braille or speech screen reader also.

The "audio/tactile" viewer has the disadvantage that a tactile copy must be printed before the figure can be read. An external digitizing pad is also required. Many kinds of information can be displayed by simpler schemes in which some kind of on-line tactile or audio object locator gives the user qualitative or semi-quantitative information about the position and shape of objects in the model. Then the user can choose to display the label for more information. We note that a three-dimensional object locator that a blind user can use to follow an object would open the possibility of access to nearly any VRML model, not just 2D projections. Finally, there are classes of information (e.g. structured trees, flow diagrams, charts, and tables) for which a user may find it more convenient to explore logical structure rather than physical structure.

These are all possible in principle, but a considerable amount of research and testing is required to learn how to translate these concepts into useful products. The special accessible VRML format we have developed permits this type of research.

The purpose of this research project is to make 2D graphics accessible to people with print impairments, not specifically to make VRML accessible. However VRML is presently the only public format that is both well-structured and flexible enough to permit addition of labels. VRML models of the future could, in principle, be made fully accessible by the authors if they choose to add sufficiently detailed labels. Alternatively, it is relatively straightforward for an editor later to add labels to a well-designed existing VRML model. To our knowledge no other 2D or 3D graphics format permits either possibility. These two properties, along with the existence of a number of user-specific special viewers make the VRML format potentially quite accessible.

ACKNOWLEDGEMENTS

This research was supported in part by the National Science Foundation under grants HRD9452881 and HRD9353094.

REFERENCES

Barry, W. A, Gardner, John A, and R. Lundquist
1994, "Books for Blind Scientists: The Technology Requirements of Accessibility", Information Technology and Disabilities, 1(4), 1994. Also available at http://www.rit.edu/~easi/itd/itdv01n4/article8.html

Bulatov, V.
(1998) Several examples of accessible VRML figures are given in the discussion at

http://dots.physics.orst.edu/vrml/
Carey, Rikk and Bell, Gavin
(1997) "The Annotated VRML 2.0 Reference Manual" (ISBN: 0-201-41974-2)
Gardner, John A.
1993 "DotsPlus - better than braille?", John A. Gardner, published in Proceedings of the 1993 International Conference on Technology and Persons with Disabilities, Los Angeles, CA, March, 1993. Also available at http://dots.physics.orst.edu/publications/csun93.html
Gardner, John A.
1998 "DotsPlus" http://dots.physics.orst.edu/dotsplus.html
Roehl, B.
(1996). VRML: A Standard for Virtual Worlds. Virtual Reality Special Report, 3(3), pp. 14-18.
VRML97 (ISO/IEC 14772)
International Standard - http://www.vrml.org/Specifications/

Biographical sketch:

John Gardner directs research programs in both Materials Physics and Information Access technologies. The latter was undertaken after losing his sight in 1988. His research is funded by the US Department of Energy and the National Science Foundation. He has won a number of prizes for excellence in both research fields.

PRESENTING DYNAMIC TACTILE BITMAP GRAPHICS USING A STANDARD BRAILLE CHARACTER GLYPH DISPLAY

M, Schönhacker [1] *)

Abstract

The author presents a simple, software-only solution to the presentation of low resolution tactile bitmap graphics on standard Braille character glyph displays. The method can be used to display static and dynamic graphics, as well as for the generation of small pictures by blind users. Due to a separated screen layout optimised for simultaneous presentation, easy cooperation of a blind and a sighted person (e.g. in a two-person game) and supervision (e.g. in a teaching environment) is possible. Sample applications shown in this paper include a simple bitmap editor as well as a mathematical function plotter.

1. INTRODUCTION

"It is completely obvious that a conventional one-line Braille display is not suited to furnish a tactile representation of a screen in graphics mode." – This quote from [4] seems natural, yet it is disappointing.

Some types of screen output can be represented in other ways. Interesting developments include a way of adapting Graphical User Interfaces (GUIs) for use by visually impaired persons [2], a method of programming GUI access [5], translation of elements of the MS Windows™ environment to speech output [1], tactile feedback of mouse actions [3], parametric sound beams [6], and many other ways of transforming screen content into media other than visual output.

This is feasible in cases where it is in fact textual content or hierarchical structure that needs to be presented. However, there are situations where tactile information is indispensable, e.g. when trying

---

[1] Martin Schönhacker, Institute of Computer Graphics, Vienna University of Technology,
Karlsplatz 13/186-1, A-1040 Vienna, Austria; Telephone: +43 1 58801 4079, Fax: +43 1 504 25 83,
Email: schoenhacker@eiunix.tuwien.ac.at

to convey knowledge of symbols a visually impaired person might encounter in particular situations. A tactile map may include simple glyphs like the common exit symbol (an arrow originating in the centre of a square), and it may be useful to learn about the glyphs commonly used in (printed) Latin script to be able to decipher engraved characters.

Thinking of more complex applications, learning about mathematical functions could be made easier by an interactive tool offering a graphical display, and blind users may wish to use their computers to play graphical games very much like sighted people do, including interactive dynamic graphics.

Unfortunately, tactile graphical displays are very expensive, so a bitmap display of reasonable size is currently unfeasible. However, a considerable number of 40– or 80–character Braille displays for Personal Computers is already in use. This article presents a simple software solution to using this kind of display for graphical information, and for displaying mixtures of graphics and Braille text.

2. A different look at Braille characters

Usually Braille line displays are being looked at as just that: a set of consecutive Braille characters. However, as each character consists of six or (more usually on computer displays) eight dots, an array of characters may be seen as a tactile graphical bitmap. For example, *Figure 1* shows a pattern that might be read as "eieie", but could also be interpreted as the graphical representation of a wavy line.

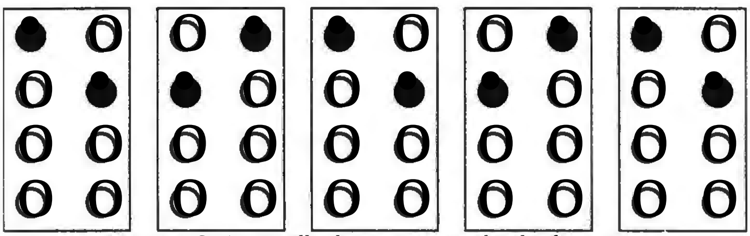


Figure 1: Seeing Braille characters as graphical information.

At first glance, using a Braille line for this kind of display appears to be a rather ridiculous and pointless effort. After all, a 40–character display will only result in a bitmap of 80 (horizontal) by 4 (vertical) dots, which is a rather strange resolution for common GUIs. On the other hand, as can be seen from the "wavy line" example in *Figure 1*, resolution is not always vital. Furthermore, scrolling can be used to display a larger "virtual" bitmap that can be scanned by the user.

The actual process of displaying graphical information can easily be programmed once a "reverse mapping" from Braille characters to individual dot combinations has been established. This list of 64 (for 6–dot displays) or 256 (for 8–dot displays) codes has to be adapted to the local or machine-specific version of the Braille character set, which basically involves a simple set-up process and needs to be done only once. Software can then take care of the rest.

In particular, dots cannot be set individually as on conventional bitmap displays. The state of all the other dots in the same Braille character must be taken into account, and a reverse mapped character code is calculated from the state of all its dots. This operation can be encapsulated in a driver module, so the actual program simply uses bitmap operations that are translated accordingly.

Concurrent display of graphics and text is yet another problem, as there is no way to distinguish a Braille character from a small array of graphical dots. One needs an additional attribute to be able to decide what is being presented, and the use of colour attributes seems to be a logical solution. Most Braille lines will provide this information on demand, be it as speech or on a separate tactile display, and by virtue of this additional feature graphics and text can be mixed. Note that this can only be done on a character by character basis, i.e. in horizontal steps of two dots per unit. This is is a serious restriction, but one that needs to be understood.

3 Taking care of the sighted user

Strange as it may seem, the proposed method of displaying tactile graphics is a real problem to *sighted* users. While a blind user is examining the graphical display, the view for a sighted person (e.g. a collaborator, partner or teacher) is impaired by the blind person's hands. Also, some Braille lines feature black dots on black background, which makes them virtually impossible to read visually.

Unfortunately, looking at the screen is not an option for the sighted user, as the character codes displayed by the reverse mapping procedure do not bear any visual resemblance of the graphical representation. Therefore a sighted user will not usually be able to understand what is being displayed on the Braille line.

The proposed solution to this problem is a screen layout that satisfies the needs of both the visually impaired and the sighted user. It involves using a part of the screen to display the (visually meaningless) reverse mapped characters for the Braille line (*Braille display area*), separating this area clearly from everything else to prevent confusion for the blind user, and displaying the same information visually on the rest of the screen. However, the matter is not quite as simple.

Firstly, textual information should of course be presented to the sighted user as text. This suggests the use of a *text display area* which only extracts the text from the tactile information.

Secondly, the bitmapped tactile display needs to be mapped to graphics on the screen. Unfortunately the Braille line will only display properly if the screen is in text mode, so the "graphic" representation has to be done in a *graphical display area* using the block graphics characters of the standard PC character set. This requires a software driver which is actually very similar to the one displaying graphics on the Braille line, as it also needs to combine two "dots" (2 rows by 1 column) into one reverse mapped character. To enhance presentation, characters on the Braille line can be shown in alternating levels of grey. Although this information is normally redundant, it proves particularly useful e.g. when a sighted person explains the display to a blind user.

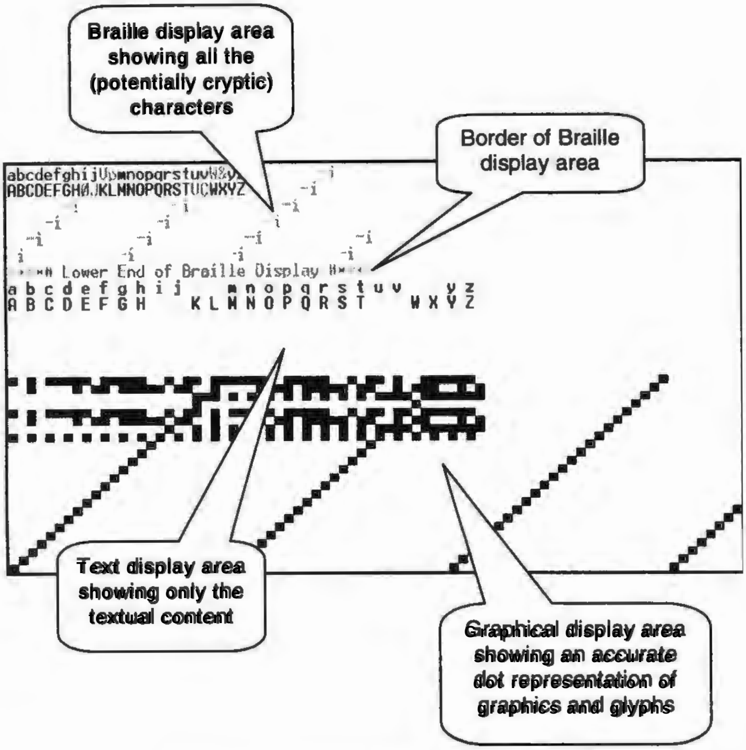


Figure 2: Experimental combined screen layout.

Figure 2 shows the screen layout used by the author in various experiments. It has been found to quite useful and easily understandable. For the purposes of this paper, the picture has been inv for better printing. Note the levels of grey indicating the difference between textual and gra information in the Braille display area. In particular, some of the characters of the alpha sequences have been overwritten and turned into graphics by the lines that can be seen in graphical display area.

The text display area has been spaced in order to correspond to the graphical display area, corresponding characters are shown in identical colours. This makes it unnecessary for the sig used to be able to (visually) read Braille script from the graphical display area.

4. Sample application: a graphics editor

Figure 3 shows the opening screen of a simple graphics editor written by the author. That screen example of text-only information. As can be seen, text is of course visible in the Braille d area as well as in the text display area. It is the inclusion of graphics that makes a separate display area so useful, because the Braille display area will then contain confusing character seen in *Figure 2*.

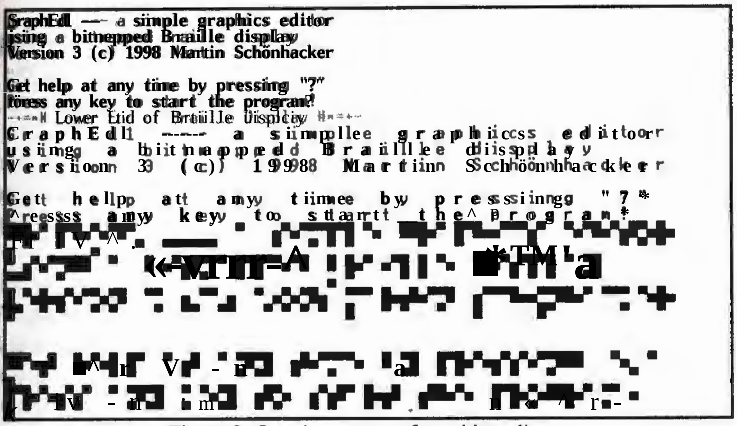


Figure 3: Opening screen of graphics editor.

The graphics editor has been used experimentally by a blind student who was almost immediately able to draw her own small pictures of geometric shapes. The program quickly turned into a favourite toy, which seems understandable: It offers a facility of generating tactile graphical information without having to spend a long time on handling "hardware", e.g. attaching threads of wool to cardboard.

When switching to "graphics mode", the experimental editor showed another small problem: It would have been useful to have a permanent cursor, but this could not be done reasonably in the graphic representation.

Consequently, a "dynamic cursor" was introduced. A special key (currently it is the Space key for convenience) makes the graphical cursor blink several times by inverting the respective dot at an experimentally determined frequency of 2½ Hz to give the moving parts in the Braille line enough time to react to the stimulus. The hardware screen cursor will at the same time be set to the corresponding line and character in the Braille display area, which should usually set the Braille line to show that place.

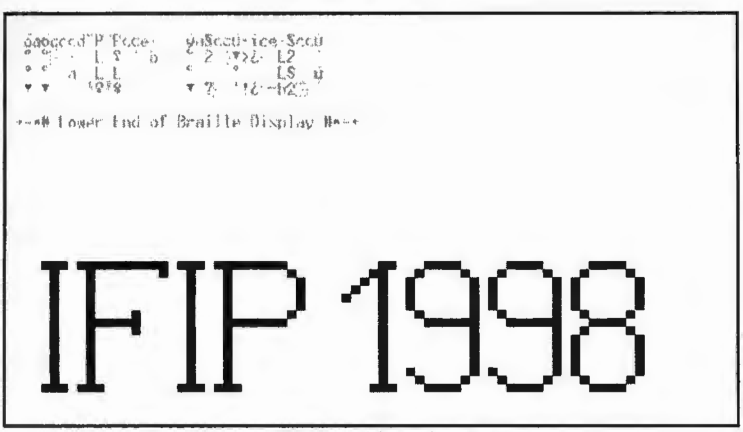


Figure 4: The graphics editor at work.

The editor can be used to draw bitmapped pictures sized 80 (horizontal) by 24 (vertical) dots as shown in *Figure 4*. Note that in this case the text display area is inactive because the entire screen area is being used for graphics. The Braille display area shows only strange characters, which can of course be expected.

In the current screen layout, the Braille display area is limited to six lines in order to make everything fit on one screen. However, horizontal and vertical scrolling can easily be programmed, which would allow a scrollable bitmap to be of almost arbitrary "virtual" size.

Future experiments will include displaying longer texts with embedded graphics. This extension brings up an interesting problem as text may only be scrolled four pixels (i.e. one text line) at a time to remain readable, while graphics could in principle be scrolled smoothly (i.e. pixel by pixel) although the usefulness of such fine-grain scrolling appears somewhat questionable.

5. Sample application: a function plotter

Figure 5 is an example of another sample application. It shows the plot of a sine wave, including labels for the axes and the function itself.

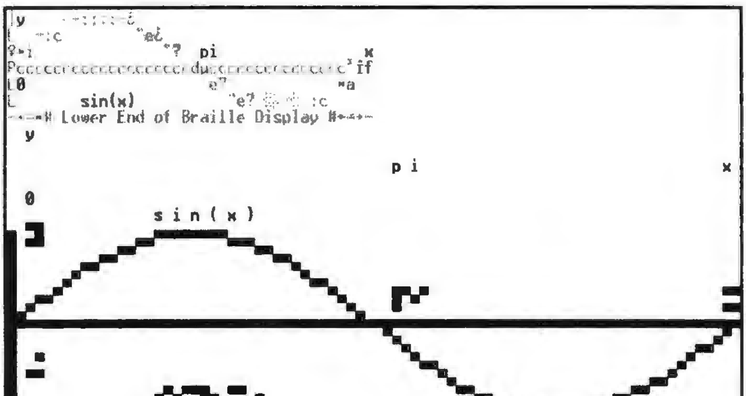


Figure 5: Plot of sine wave including descriptive text.

As can be seen from the success of various pocket calculators capable of plotting mathematical functions, graphical resolution is not the main requirement for this type of application. And if it can be done using a standard device like a Braille line, a function plotter could certainly enhance a visually impaired person's access to some aspects of mathematics that might otherwise be hard to understand.

6. Future perspectives and acknowledgements

There are many applications of this simple tactile graphical display, in particular considering that it does not generate any additional cost in hardware if the visually impaired person's computer is equipped with a Braille line.

Initial experiments with dynamic display have already been conducted, and the results are promising. For instance, a few dots moving along the display can be declared to represent asteroids which brings a classic computer game to the world of tactile graphics. As the information is also visible on the screen, the logical next step would involve computer games for two users, where one is sighted and uses the screen while the other is visually impaired and uses the Braille line.

Research on this project has partially been made possible by a grant from the "Hochschul-jubiläumsstiftung der Stadt Wien", for which the author wishes to express his thanks and gratitude.

7. REFERENCES

[1] GRAZIANI, P. and BRESCHI, B., Screen Reader for Windows Based on Speech Output, in: Computers for Handicapped Persons (eds. Zagler, W.L., Busby, G., Wagner, R.L.), Proceedings of 4th International Conference, ICCHP '94, Vienna, Austria, 1994.

[2] KASHMER, A.I., OGDEN, B., BRAWNER, P., KANGAS, K. and REISWIG, G., Adapting Graphical User Interfaces for Use by Visually Handicapped Computer Users: Current Results and Continuing Research, in: Computers for Handicapped Persons (eds. Zagler, W.L., Busby, G., Wagner, R.L.), Proceedings of 4th International Conference, ICCHP '94, Vienna, Austria, 1994.

[3] KERSTNER, W., PIGEL, G. and TSCHELIGI, M., The FeelMouse: Making Computer Screens Feelable, in: Computers for Handicapped Persons (eds. Zagler, W.L., Busby, G., Wagner, R.L.), Proceedings of 4th International Conference, ICCHP '94, Vienna, Austria, 1994.

[4] STÖGER, B. and MIESENBERGER, K., The Conventional Braille Display – State of the Art and Future Perspectives, in: Computers for Handicapped Persons (eds. Zagler, W.L., Busby, G., Wagner, R.L.), Proceedings of 4th International Conference, ICCHP '94, Vienna, Austria, 1994.

[5] THATCHER, J., Screen Reader/2 – Programmed Access to the GUI, in: Computers for Handicapped Persons (eds. Zagler, W.L., Busby, G., Wagner, R.L.), Proceedings of 4th International Conference, ICCHP '94, Vienna, Austria, 1994.

[6] YONEZAWA, Y., NONAKA, H., ITOH, K. and HASHIMOTO, M., Projective Display of Document Information by Parametric Sound Beam, in: Computers for Handicapped Persons (eds. Zagler, W.L., Busby, G., Wagner, R.L.), Proceedings of 4th International Conference, ICCHP '94, Vienna, Austria, 1994.

Session IV

Access to the World Wide Web

WebAdapter: A prototype of a WWW browser with new special needs adaptations

Dirk Hermsdorf[1]

Abstract

This paper will present the prototype of the WWW (World Wide Web)-browser WebAdapter, which provides new special needs adaptations for physically handicapped, blind and visually impaired end users. These adaptations include near miss tolerances, implementation of sophisticated HTML-guidelines and advanced speech output. A user test will evaluate the usefulness of these features. Other improvements for the user interface of WWW-browsers will be discussed as well.

1. Introduction

Computer technology has become a very important part of our all-day life. It can be found in offices, factories, our homes and many other places. Computers are a necessary tool to handle the mass of existing data. So it is very important that everyone including people with special needs can access them otherwise some people will not have access to certain informations. This problem can be solved by designing a user interface for all (UI for All) [19].

Especially for people with special needs the Internet and in particular the WWW is a very important source of informations. It provides a lot of different, widespread informations in a familiar and ubiquitous environment.

Informations in the WWW are obtained by so called WWW-browsers. Recently new versions of common WWW-browsers like Microsoft Internet Explorer 4.01 and Netscape Navigator 4.0 try to improve their user interface (UI) for people with special needs. Here one of the most important improvements is a complete access by keyboard for the benefit of physically impaired users, who do not need to use the mouse anymore. The WebAdapter provides other, new improvements of the user interface, which are not integrated in any other WWW-browser so far.

The WebAdapter was developed within the scope of GMD' s project TEDIS (TEleworking for Disabled people) [9].

---

[1] German National Research Center for Information Technology (GMD), 53754 Sankt Augustin, Phone +49 2241 14 2018, Fax +49 2241 14 2065, Email dirk.hermsdorf@gmd.de

2. The WebAdapter

The WebAdapter provides common functionality like storing bookmarks, a context sen function and additional adaptations for blind, visually impaired and physically handica The intention of this prototype is to illustrate the possibility to include adaptations for special needs in a world wide web browser without affecting the UI of non - handicapped users.

2.1 Near miss tolerance

The near miss tolerance, which is integrated in the WebAdapter, enables the user to select but the symbol bar even when the mouse pointer is close but not directly touching the button for of physically impaired users. The closest button is always highlighted when this feature is serving as a visual feedback so that the user knows which button is selected when the right button is hit.

The near miss tolerances are so far only implemented for the symbol bar. An extension for links, maps and so on would be useful, but could not be implemented due to interface problems with programming environment (see 2.5).

2.2 Automatic application of HTML guidelines

An other approach realises the implementation of some HTML (Hypertext Markup La guidelines [2, 3, 4, 5, 7, 14, 15, 16, 17, 18, 20, 22]. These guidelines provide specific insin web designers to program a more accessible web side for people with special needs. Applyi guidelines should nevertheless not affect the aesthetics of the HTML-document so th handicapped users are not disturbed by the adaptations. Some of the HTML-guidelin wh implemented in the WebAdapter can be applied automatically. For example one of these gu demands to label lists (largeness of list, number of sublists) so that blind users can get t overview about the document structure, here the list structure.

A lot of the implemented guidelines realize a better presentation for screen readers. Screen re software programs, which read the textual screen contents to the user.

Altogether seven HTML-guidelines which could be implemented in that way were found. Th specify in the WebAdapter which adaptations should be applied to the current HTML-do a technical point of view, a guideline is applied by parsing and changing the HTML-co document.
The following adaptations are provided by the WebAdapter:

1 adaptation for physically impaired users:

- It is possible to specify a minimum size for clickable images. This is good for physically users, who can enlarge the image and hit it more easily.

1 adaptation for visually impaired users:

- Turn off background image.

5 adaptations for blind users:

- Replace acronyms. Some screen readers have problems with letters like $, &, % and so on. This feature enables the user to replace acronyms by other expressions like dollar and percent and so on (also look at the ACRONYM tag in HTML 4.0 described in chapter 5).

- Label lists.

- Insert dots in abbreviations like LSU, GMD etc. so that a screen reader will pause between each letter and will not try to read the abbreviation as one word. An abbreviation is defined as two capitals following each other immediately. A wrong interpretation of abbreviations can not be excluded - for example the letters in a header might be all in capitals.

- Insert a dot at the end of list elements, so that the screen reader will lower its voice at the end of each list item pointing out its end.

- Sequential presentation of tables (also look at HTML 4.0 style sheets described in chapter 5).

The last approach will need some detailed explanation. A lot of HTML-guidelines focus on the accessibility of tables because they are difficult to handle. Many of these guidelines recommend to abandon the use of tables at all. But this will not be possible because tables provide a too powerfull tool to present a lot of informations in a compact manner. Additionally, many web designer will not obey to such a restriction, because the aesthetic of the HTML-document would be affected.

The problem with tables is, that screen readers read the presented information line by line which totally confuses the listener when tables are concerned. Additionally some screen readers do not pause between the different cell entries.

The WebAdapter provides six sequential counterparts of a X*Y table (X rows, Y columns). Providing a sequential form of a table will improve the accessibility of tables. But how is it possible to achieve a sequential form?

The problem is that there are different ways how a table can be read. A seeing user can decide just by looking at the table how to read it. This knowledge can unfortunately only be implemented partly. The WebAdapter extracts certain informations about the table form and then assigns the most appropriate sequential form. It looks for possible row or column headers which might be indicated by a bold font or the use of the <TH> tag and the distribution of different data types (integers, strings etc.). One example shows a table of a table tennis competition between Mark and Dirk:

Mark	Dirk
21	13
9	21
19	21

A screen reader would most likely read this table as "Mark Dirk 21 13 9 21 19 21" which ma
much sense.

The WebAdapter on the other hand would recognise the distribution of the different data typ
assume that the first row contains headers for the two columns. For that reason, the most appr
sequential form presented by the WebAdapter would be the following:

Mark: 21.
Dirk: 13.
Mark: 9.
Dirk: 21.
Mark: 19.
Dirk 21.

The WebAdapter can propose the most appropriate sequential form of the table or just provide a
sequential forms.

2.3 The project WAB (Web Access for the Blind)

The approach of modifying HTML-layout has already been realised in the WAB (Web Access fo
Blind [21]) project of the university of Zuerich. If you specify the proxy serv
"http://ea.ethz.ch:8080" in a common web browser, the HTML layout will change accordin
three following adaptations. Nevertheless WAB provides no modular selection of the adaptati
blocks the selection of a proxy. The modular selection of the adaptations is implemented
WebAdapter. The following adaptations are concerned:

- Provides a list of all links in the HTML-document. Seeing users can just get a quick overvie
document by scrolling through it. Blind users do not have this ability. This feature will never
enable blind users to get a quick overview or impression of the document.

- Provides a list of all headers in an HTML-document. It is possible to click on the headers. Th
the user jumps to the corresponding header in the document.

- Keywords like 'Image', 'Radiobutton', 'Checkbox', 'Editfield', 'Button', 'Combobox', 'ref'
stands for reference, that means a link that points at a position within the same web docum
'Link' are inserted in the corresponding places in the web document. So the user is infor
the object type when he is using a screen reader or braille display. Furthermore, he can sea
keywords which allows him e.g. to jump from link to link.

14 Integration of a screen reader on HTML-level

Ueijnthetic speech output Speak&Win from ETeX, Frankfurt is integrated in the WebAdapter. The BiKi/nthetic speech is integrated in the pwWebSpeak browser from Productivity Works [15].

losing the integrated speech synthetic, the WebAdapter can read an HTML-document on HTML-level. fJjis reveals an advantage in contrast to ordinary screen readers, which only can read the screen contents. The process of visualising HTML-code turns explicit structure information which are feeiified in the HTML-code into implicit structure informations. It is hard and often even impossible to »*discover these structure informations [8].

Tfaiteraction of the user with the WebAdapter can be commented acoustically. With this option it is possible to interact with the WebAdapter when the monitor is turned off. The only input device can be lie keyboard. Almost all objects are assigned to shortcuts. Additionally the user can tab through all (iijeflfwithin an active window of the WebAdapter interface. This should make the system accessible fiit blind users.

Mioreover it is possible to assign a pitch, speed and volume value to each HTML-tag. So the structure rfibcHTML-document can be presented acoustically.

During the reading process, the last read link can be selected by pressing "a" ("a" for anchor). OptiBally the system reads the current link and asks if this link really should be selected.

Altitionally it is possible to jump between the different HTML structure elements. Therefore the keys 'o' (read the current element once again), "b" (backward with a specified step range) and "f' (forward Mit a specified step range) are provided. The step range in combination with the keys "b" and "f' can be^ified by the user. A step range can be a fix jump value, words, sentences, paragraphs, links, hries, list elements and table elements. This selection can be performed in the general preferences of ibiWebAdapter or with the key "s" (set step range). Within tables it is possible to navigate with "r" (pht), "l" (left), "u" (up), "d" (down), "h" (row header) and "c" (column header) between the different o n cells. Optionally the current row and column number is spoken as well, which is only changed in iffibination with the keys "r", "l", "u" and "d". The acoustic navigation in tables starts in the first ctfmi and the first row when one of these keys is pressed and can only be aborted with the key "e" rt^Then the speech output jumps to the end of table. The acoustic navigation in hypertext systems is implemented in a similar way in other research projects [1, 13], but without a table navigation. [1] implemented additionally a fast forward function. The integration of a fast-forward and rewind function might be a useful extension of the current system.

The adaptations do not need to be visualised. This shows the "dual use" aspect of the WebAdapter. While a sighted impaired person looks at the unchanged version of the current web document, the blind jBSon who sits at the same computer listens to the adapted version of the document (using headphones).

2.5 Implementation

The WebAdapter was implemented with Visual Basic 4.0 and the Internet Control Pack [6], bec was the only possible environment for a quick implementation. Since 1. April 1998 the source o Netscape's Communicator is available at www.mozilla.org for everyone. An integ WebAdapter's features for people with special needs into a common WWW-browser would possible and would have avoided the implementation of a prototype.
The beta version of the Internet Control Pack is distributed by NetMasters. The Internet Con consists of 8 OLE-controls (object linking and embedding), which provides useful functional creating Internet applications. Only one of these controls, the HTML-control, was used to prog browser.

The HTML-control supports the following features:

- Scrolling of the selected page.
- Inline graphics: GIF, JPEG, BMP, XBM.
- HTML version 2.x and most of the Netscape 2.0 and Internet Explorer 2.0 extensions.
- Integrated document retrieval for http and file URLs.
- Integrated http extension for form elements.
- Properties for determining the document form (like fonts, colors)
- DocInput, DocOutput object interface for flexible data exchange.
- Events to overwrite default processes.

3. Evaluation of the WebAdapter

The browser will be evaluated by usability tests and questionnaires. A corresponding in guideline basing on the ISO 9241 norm will be developed.

One evaluation will regard the adaptations for physically handicapped users. Therefore a pi task will be performed by two physically handicapped adults. A possible task might be to fi author's homepage starting from the GMD homepage www.gmd.de.
The other evaluation will evaluate the usability of the keyboard as the only input device in con which speech output as the only output device. Visually impaired users as well as non-ha users will perform the predefined task.

4. Other improvements for web browser user interfaces

This list proposes further improvements of the user interface of WWW-browsers [10, 11, 12, are not implemented in common WWW-browsers like Microsoft Internet Explorer 4.01 and Navigator 4.0:

ternative format for information objects:

- Presentation of the description field of PNG graphics.
- Presentation of text or static graphics instead of java applets.
- Presentation of the ALT text in client-side imagemaps.

Tables:

- Navigation with arrow keys through tables.
- Presentation of textual information of tables (AXES and AXIS attributes) in order to distinguish the current row and column.

Frames:

- Navigation through frame objects.
- A frame on/off option.

Java:

- A java on/off option.
- Presentation of an alternative description for java applets (APPLET ALT).

Orientation:

- A "Where am I" option. This should provide informations of the current position of the user within the document like "I'm 34% through the entire document, I'm in the second frame, the third header, the fifth list element of a ten element list ...).
- A page summary like "this side contains X images, Y frames, Z headers, ..".

Navigation:

- Navigation per keyboard through links, headers, paragraphs, ...
- Determination of a special keyboard command in order to tab through links in a selectable step range.

Feedback:

- Visual and acoustical feedback for user interactions and occurring events (commented interacts the WebAdapter).

- Specification of the font size, color and button size.

Presentation:

- A high contrast mode (font sans serif, background white, normal text black, visited links vi other links blue).

- A zoom modus.

Interaction:

- Recognition of words after only some letters have been typed (definition of macros, dictio Netscape 4.0 and Internet Explorer 4.0 provide this feature in the Open URL text field.

- A keyboard command to mark text.

- Selection of links by typing in the link's label first letter (like the selection of windows object

- Integration of a scanning mode. In this mode, selectable objects and object groups are highlig one after another. The highlighted object can be selected by pressing any key enabling the us choose objects with only one key.

5. Conclusion

HTML 4.0 will was released in December 1997 provides a lot of new features for the benefit of p with special needs. Nevertheless will HTML 4.0 not replace the necessity of improvements of browsers user interface, like the ones illustrated in the WebAdapter and in chapter 4.
The most important HTML 4.0 tags for people with special needs are:

- STYLE: Strong emphasis on distinction between document structure and presentation (t style sheet integration and deprecation of HTML presentation elements and attributes).

- ALT tags now required on images, to make images easier to understand when gra presentation is not used.

- LONGDESC attribute an added option for IMG, FRAME & TABLE, to enable linking to ad description of images, frames or tables.

- OPTGROUP element added within the SELECT FORM controls to make long lists of choices more manageable when they are serialized.

- FIELDSET, LEGEND gives the ability to group FORM controls semantically.

- ACCESSKEY and TABINDEX provides better navigation within forms.

- SUMMARY, CAPTION, COLGROUP, SCOPE and HEADER attributes added to make TABLE easier to understand when linearized.

- MEDIA: Wider range of target media (TTY, Braille, etc.) for use with style sheets.

- OBJECT provides the ability to furnish hypertext description of an included object.

- MAP has a new content model that allows client-side image map to provide more detailed textual explanation of the links in the map.

- TITLE and LANG support on all elements allows systematic indication of document structure and content.

- ABBR and ACRONYM elements support systematic description of acronyms, which is very useful for speech output.

The use of style sheets should replace the use of tables for layout purposes. But not every web page designer will replace his table layout. Even the most accessible HTML-standard is useless until it is applied by the web designer. But new standards like the required use of the ALT tag will definitely improve the web page accessibility.

HTML 4.0 was designed by the W3C' s HTML4 Working Group in co-operation with the Web Accessibility Initiative's (WAI) HTML & CSS Review Working Group. The GMD as a W3C member also participates in the WAI, which was launched to the occasion of the last WWW6 conference in Santa Clara, CA in April 1997.

i. References

[1] Chieko Asakawa, Takashi Itoh, "User interfaces of a home page reader", IBM Japan, in: ASSETS '98, the third international ACM conference on assistive technologies, April 15-17, 1998, Marina del Rey, CA, ACM press, ACM ISBN: 1-58113-020-1

[2] Adaptive Technology Research Center, University of Toronto, "Guide to accessible HTML, Accessibility for Persons with visual disabilities", http://www.utoronto.ca/atrc/rd/html/html.html

[3] Adaptive Technology Research Center, University of Toronto, "Making the Web Accessible for the Blind and Visually Impaired", Mike Paciello, http://www.yuri.org/webable/mp-blnax.html

[4] DO - IT (Disabilities, Opportunities, Internetworking & Technology), "DO - IT HTML Gid http://weber.u.washington.edu/~doit/Other/design.html

[5] Paul Fontaine, Center for Information Technology Accommodation, General Services Administration, Washi USA, 1995, "Writing Accessible HTML Documents", http://www.yuri.org/webable/htmlcode.html

[6] NetMasters FastNet ActiveX, http://www.startups.com/products/fna/default.html http://www.netmastersllc.com/products/fna/

[7] InfoUse, Center for Accessible Technology (CforAT), "The User Interaction Design Guidelines Data 28.12.1995, Version .955, http://www.infouse.com/disabilitydata/guidelines.html

[8] Frankie James, "AHA: Audio HTML Access", in "The sixth international world wide web conference - proce April 7-11 1997, Santa Clara Convention Center, California, USA, ISBN 0965761 http://proceedings.www6conf.org/HyperNews/get/PAPER296.html

[9] Henrike Gappa, Stefanie Mermet, Michael Pieper, "Teleworking for Disabled People - Pitfalls and Challenges", to be published at ICCHP '98

[10] Jon Gunderson, "World Wide Web Brwoser Access Recommendations", http://www.staff.uiuc.edu/~jongti(Ja browsers.html

[11] Jon Gunderson, Gregg Vanderheiden, D. Shumila et al., "List of Critical W http://trace.wisc.edu/text/guidelns/browser/suggest/browsers.html

[12] Dirk Hermsdorf, Michael Pieper, "BSCW for Disabled Teleworkers - Usability Evaluation and Interface Adap an internet - based Cooperation Environment", in: Michael Genesereth and Anna Patterson (ed.), Proceedings for its International World Wide Web Conference, Santa Clara, CA, USA, 7-11.4 1997, ISBN 09657 http://proceedings.www6conf.org/HyperNews/get/PAPER161.html

[13] Sarah Morley, Helen Petrie, Anne-Marie O'Neill, Peter McNally, "Auditory navigation in hyperspace: Desi evaluation of a non-visual hypermedia system for blind users", in: ASSETS '98, the third international ACM confe assistive technologies, April 15-17, 1998, Marina del Rey, CA, ACM press, ACM ISBN: 1-58113-020-1

[14] Jakob Nielson, SunSoft Distinguished Engineer, "Accessible Design for Users With Disabilities", Octo http://www.sun.com/columns/alertbox/9610.html

[15] The Productivity Works, The pwWebSpeak Project, http://www.prodworks.com/pwwebspk.htm

[16] SIH Laeromedel Datapedagogen, Christer Johansson, "Designing information for WWW - Tips and guideli increasing accessibility for people with disabilities", http://www.sih.se/dh/ehtmldok.htm

[17] Clas Thoren, "Nordic guidelines for computer accessibility", ISBN 91-86954-15-6

[18] Jutta Treviranus, Chris Serflek, University of Toronto, "Alternative Access to the World Wide I http://www.utoronto.ca/atrc/rd/library/papers/WWW.html

[19] Gregg C. Vanderheiden, "Anywhere, Anytime (+Anyone) Access to the Next - Generation WWW", in Th international world wide web conference - proceedings", April 7-11 1997, Santa Clara Convention Center, Californ ISBN 0-9657614-0-1, http://proceedings.www6conf.org/HyperNews/get/PAPER253.html

[20] Gregg Vanderheiden (Trace Research and Development), Wendy Chisholm (Trace Research and Developm Jacobs (W3C), "WAI Accessibility Guidelines: Page Authoring", W3C Working Draft, 14-A http://www.w3.org/WAI/GL/WD-WAI-PAGEAUTH-0414.html

[21] Projekt "Web Access for Blind Users", Institut fuer Informationssysteme, ETH http://www.inf.ethz.ch/department/IS/ea/blinds/

[22] Adlai Waksman, "Making your web pages Lynx friendly", http://www.cfar.umd.edu/~waksman/lyn

WEB BROWSING FOR THE VISUALLY IMPAIRED

M. Zajicek, C. Powell, C. Reeves, J. Griffiths

Abstract

The aim of our work is to enable visually impaired users to effectively browse the Web alone or with sighted co-workers. We have built a prototype browser which uses information retrieval to provide a set of complementary options to summarise the Web page. Synchronised visual representation enables co-work with sighted users.

1 Introduction

At the Speech Project at Oxford Brookes University we have built a Web navigation tool called BrookesTalk that employs information retrieval techniques to summarise a Web page for quick orientation [7]. This paper describes the facilities embedded in BrookesTalk. It includes a brief description of the information retrieval techniques used to extract keywords and abridged text from a Web page.

The paper also describes our evaluation of the usability of BrookesTalk particularly the keywords and summary information facilities. We discuss the future for information retrieval techniques in page orientation and then discuss some of the questions our evaluation raises.

2 The aims of the speech project

Speech output screen readers for the blind such as JAWS and browsers such as pwWebSpeak™, enable the user to move around the screen, jump to interaction objects in the screen and read particular areas of text enabling the user to build up a picture of the screen and move around within it.

These tools become unwieldy when moving quickly around the Web searching for information and making hasty decisions about which pages or bits of pages might be useful. Users are often and which pages might be useful within themes. In addition visually impaired people need to on the Web with sighted users.

The aim of BrookesTalk is to support users in this type of activity where speedy decisions about usefulness of a page or part of a page are paramount [6].

3. How BrookesTalk works

BrookesTalk is a small speech output browser which is independent of visual browsers and independent of text to speech software applications as it uses Microsoft speech technology.

It includes the functionality of a standard Web browser for the blind [4] such as pwWebSp that it can break up the text part of a Web page into headings and links and read out paragr However the main aim is to provide an orientation tool for blind users in the form of a toolbar of functions that will provide different synopses of a Web page to help the user whether it will be useful to them or not.

Users can select from a menu of list of headings, list of links, list of keywords, list of an abridged version of the page, a list of scratchpad entries, a summary of the page, and reach and read out chunks of text which are organised hierarchically under headings. It is exp that the user will pick tools from this virtual toolbar which complement one another particular type of page under review.

While the speech driven browser is our first priority, the need for blind or visually impaire to co-work with sighted users is paramount for total integration into the workforce. In ad people with residual sight need to use all senses that can help them. For this reason we included different visual representation of the Web page as a facility in BrookesTalk. Users elect to run a standard visual browser concurrently and also display a large text versio headings, links and keywords. An indicator shows which word is being spoken at any time.

The list of keywords consists of words which are assumed [1] to be particularly meaningful the text. These are found using standard information retrieval techniques [5] based on frequency. Abridged text is also compiled of special sentences which have been isolated using trigrams.

The scratchpad allows users to save any sentence they are listening to, which they con important or worth noting, simply by pressing a key. They can then playback lists of sen linked to particular pages. The summary of the page includes author defined keywords, the number of words in a page, the number of headings and the number of links.

4 Keywords and abridged text

Document title	Total no.of words	Keywords	Document subject
Thinking of getting a border collie?	1174	border, collie, people, collies, dog, exercise, dogs	Information for prospective border collie owners
Comet Hale-Bopp Home Page (JPL)	772	comet, hale, image, January, images, vodniza,	Information about comet Hale-Bopp
Dennis' Information Information Retrieval Page	279	information, indexing, IR, pages, index, search, engines,	General about information retrieval and links to related sites

Table 1 Examples of Keyword Extractions

A list of keywords for three different Web pages are shown in Table 1.

It can be seen that the small number of keywords could be used to augment or replace information given in the title if it was not available or perhaps not informative.

When keywords are used contextual information imparted by the position of a word in a sentence is lost. Extraction of three word key phrases or trigrams preserves some word position information. The technique is based on 'Word level n-gram analysis' in automatic document summarisation [3] To provide a measure of similarity, groups of words appearing together rather than individual words are compared.

To reduce the number of word-level mismatches due to the normal changes in spelling required by grammar, each element of a trigram was assigned the stem of a word [2] rather than the word itself. The trigrams presented were ranked by frequency.

Trigram	Frequency	Content Words	Keywords	Score
equal opportun polici	2	3	3	8
equal opportun and	2	2	2	6
servic deliveri and	2	2	0	4
it is commit	2	1	0	3
is commit to	2	1	0	3

Table 2 Scores for Trigrams and their components on the RNIB 'Equal Opportunities Policy' page

High frequency trigrams occur twice, low frequency trigrams once, providing at first glance little distinguish between them. Many of the words in the trigrams are noise words which are require for grammatical correctness and are not content bearing. A summation of frequency of trig number of content words in the trigram and number of keywords in the trigram appears as th score for the trigram in Table 2. which shows key trigrams for the RNIB 'Equal Opportunities Policy' page.

Abridged pages were created by computing the key trigrams of a page, according to score, and th creating a page consisting of the sentences in which the trigrams appeared. Abridged pages i average worked out to be 20% of the size of the original text and, unlike keyword lists, composed of well formed (comprehensible) sentences.

5. Evaluation of BrookesTalk

5.1 Keyword evaluation

Preliminary experiments were performed to assess the usefulness of the keyword list as an inde of page content compared to the headings list or the links list. Headings and keywords in partic were judged to be roughly comparable in that they provided a list of indicating words or phrases.

The argument for incorporating keywords in the BrookesTalk menubar is that it provides more flexibility for the user in summarising the Web page. If the author has truly encapsulated the meaning of subsections of the page in headings then headings should provide a significantly better indicator of page content than keywords. However headings are often represented as images whieh do not provide speech output, or are eye-catching rather than informative. In this case keywords could provide a better summary. The aim of BrookesTalk is to provide a range of tools to aid orientation thus overriding many of the vagaries of Web page authoring.

Users' perception of the usefulness of the representation was measured by asking them to evaluate the usefulness of describing a Web page using the three different types of summary representations, headings, links/anchors, and keywords.

Twenty subject users were given the different representations for six different Web pages. The pages were chosen to maximise variability. Subjects gave a seore between 0 and 5 for each representation. The sum of the scores for each representation, together with the percentage of the total score it represented, was taken to give an indication of its effectiveness. Results are shown in Table 3.

We see that users perceived that keywords provide a considerable improvement on the use of links to orientate users to Web pages. Headings gave the best score but the score for keywords was not significantly different.

	Total Score	Percentage of available marks
Headings	406	58.09
Anchors	313	45.24
Keywords	392	55.28

Table 3. Scores for different representations

5.2 Experiments with summary information

The page summary was perceived to be an important tool in Web page orientation. pwWebSpeak(TM) a leading browser for the blind gives the numbers of headings, links, and images as its summary. The BrookesTalk summary comprising title, author name (if any), author defined keywords (if any) number of words in the page, headings, links and extracted keywords was found to be more useful. The number of words in a page was found to be particularly useful for page orientation.

Special tests were run to determine the relative value of author defined keywords compared with extracted keywords. The vast majority of Web page authors do not include author defined keywords. The main purpose in including them being to facilitate inclusion in searches. When author defined keywords did occur they were more valuable for page orientation than extracted keywords

5.3 Usability and the BrookesTalk environment

The prototype BrookesTalk was used by a group of blind users including those at the Royal National Institute for the Blind (RNIB). User acceptance was no problem as this group was committed to finding out what software is available for blind people. They were all technically able, although our ultimate goal is to develop software for non-technical users so that all blind people can use the Web.

Earlier versions of BrookesTalk required TextAssist software for the speech synthesis. This software required patching in Windows'95 and caused discouraging technical complications before getting started. BrookesTalk was then re-written to run on the Microsoft Speech engine for Windows. While increasing portability users were concerned that the speech engine currently uses a lot of disk space. Both versions are available.

BrookesTalk uses different voices for conceptually different parts of a Web page. This is appreciated by most but described as irritating by one. We plan to make different voices optional in the future.

Users felt that it was an advantage for sighted co-workers to work with visually impaired users so that it could often be useful when clarification was needed. Synchronisation of visual displays and speech was a problem as speech output 'jumps' rapidly around the page. The sighted user must scroll the page independently. Some speech output has no equivalent screen pattern. Users commented that the speech output of headings and links also helped sighted users to orientate themselves to a page.

5.4 Evaluating the BrookesTalk virtual toolbar

Users were observed to rely heavily on one function rather than move between different summarising representations. They had been encouraged to try using the different functions know what type of page they are searching for, research work, entertainment, product details etc. They therefore know how useful headings are likely to be and can use keywords accordingly.

Surprisingly one user orientated himself by using the 'movement between links' key 90% of the time. We had not anticipated that he would build his conceptual model of the page by looking at the links behind it. This approach will be investigated fully!

The abridged version of the page received most criticism. The trigram analysis could easily pick out the wrong trigrams as being significant and important headings were frequently left out of the summary. The algorithm for picking trigrams is not very stable it can easily be influenced by irrelevant words. It was suggested that trigrams should carry some kind of semantic weighting if they appear in the title or headings.

The scratchpad worked well and provided an easy way of saving important sentences from the page. As yet sentences are linked to pages and a new scratchpad must be started with each new page. Users suggested that sentences could be tagged as related to search themes. In this way sentences from several different pages could be grouped by theme.

The role of information retrieval techniques

We have seen that keyword and trigram spotting has contributed to page summarisation, but how far can these techniques be exploited? They are purely statistical and rely on counting words to determine their importance. This is surely rather a clumsy representation of an entity as complex as understanding.

Information Retrieval technology was developed to aid the location of documents in a library or similar archives that match an example document, i.e. documents that are related to the content of the example document. This works well in specific areas that have their own sets of descriptive words, for example biochemistry.

The world wide Web contains a wide range of unrelated, un-refereed, diverse and freely composed documents. The content of any document is not under strict control as is the case for, say, journal articles, and so presents a less than ideal source for the information retrieval techniques to work upon. Using information retrieval techniques to aid in Web page orientation requires that the Web page conform to the expectations of those techniques. This is not always the case. Furthermore we make the assumption that a document is a complete entity dealing with one topic. We plan to extend our page summarisation to include an analysis of the page as a multi-subject document.

6.1 Multisubject documents

The method of determining keywords described above fails when the document under examin contains more than one subject. This is because the words from each subject tend to dilute ea other, leading to a mixture of unrelated keywords being generated. A partial solution to t problem can be obtained by changing the focus of keyword generation from the whole document individual paragraphs. The sentences comprising a paragraph are traditionally closely related. The identification of sets of keywords relating to the different subjects contained within a docum will aid the user in the orientation process. The number of subjects in a document may also b useful as an indicator of the difficulty in locating any particular piece of information. Prelimin paper studies have indicated that the procedure is worth investigating further, and software capab of multiple subject determination is being developed.

6.2 The limits of information retrieval

The problem with using information retrieval technology to provide orientation clues is that depends entirely on the statistical rules it embodies. The richness of any language allows a vari of words to be used to convey the meaning. The statistical methods employed here would fail t notice the connection between two documents using disjoint word sets to describe the same obje or event.

In order to obtain relevant keywords and subject areas in these instances, the underlying meaning of the document and its paragraphs must be taken into account. In other words the system mu *understand* what the document is about. The structure of knowledge bases and analysis metho capable of simulating the understanding of text are currently being researched in The Intellig Systems Research Group at Oxford Brookes University and it is intended to optimise BrookesTa using tools developed by the group.

7. Conclusion

Initial evaluation highlighted the potential usefulness of features designed to improve navigation, such as the use of keywords and page summary. Blind users emphasised the potential of a tool sueh as BrookesTalk to sort through what they referred to as the 'increasing pile of paper that arrived on their desks' during the working day. Methods used to translate HTML formatting to speech can be applied to other formatted documents.

Evaluation has shown that users have many different approaches to the use of BrookesTalk. Some stick with a small set of options and use them all the time. Others make what at first appear to be quite perverse choices of facilities, for example the user who orientated himself to the Web using links only. How do we perceive a Web page? Do visually impaired users perceive it differently from sighted users? The British Computer Association for the Blind (bcab) mailing list has recently dealt with the ambivalent attitude of the visually impaired towards visual images. We continue our work to determine the best way of encapsulating the information in a Web page.

8. References

[1] Luhn H P. The automatic creation of literature abstracts. IBM Journal of Research and Development, 2, 159-165. 1958

[2] Porter, M. F. An algorithm for suffix stripping. Program Automated Library and Information Systems, 14 (3), 130-137. 1980.

[3] Rose, T., and Wyard, R. A Similarity-Based Agent for Internet Searching. in Proceedings of RIAO'97. 1997.

[4] Zajicek M., Powell C., Enabling Visually Impaired People to Use the Internet, IEE Colloquium 'Computers helping people in the service of mankind', London. 1997.

[5] Zajicek M., Powell C., The use of information rich words and abridged language to orient users in the World Wide Web, IEE Colloquium 'Prospects for spoken language technology', London. 1997.

[6] Zajicek M., Powell C., Building a conceptual model of the World Wide Web for visually impaired users in Contemporary Ergonomics 1997 S. A. Robertson (Ed). 1997.

[7] Zajicek M., Powell C., Reeves C., Orientation of Blind Users on The World Wide Web in Contemporary Ergonomics 1998. 1998.

AN INTERNET BROWSER FOR VISUALLY HANDICAPPED USERS: PRINCIPLES AND METHODS

Djamel HADJADJ, Abdellah BOUZIDI, Dominique BURGER

Abstract

This paper describes the development of a non-visual Web browser as part of the Fren BrailleNet[1] project, "Internet for the Education of the Visually Handicapped". A general softw architecture is first described which is compatible with development environments such a Microsoft ActiveX ™ . The software has two main functions 1) processing the HTML source to make the final HTML document more accessible to the visually handicapped users and 2) provi a user interface which combines Braille, speech, large print output and several possible keyb to make easy access and rapid navigation. The techniques used ensure that adaptation i independent of the production and the delivery of documents. These techniques are also comp with existing recommendations for HTML design.

1. Introduction

It has always been difficult for handicapped people to gain access to sources of written inform This appears to be one of the main barriers they encounter in modem society in their social a professional integration. This is particularly true in education, where written material are ess The recent advent and dramatic growth of the Internet has completely revised this problem seems to be about to provide efficient solutions.

BrailleNet was created in September 1996 in order to promote the use of Internet in the educ of the visually handicapped in France. It includes user organisations, schools and univer research laboratories and industrial companies [1]. It main achievements during the first year of t project have been the creation of an educational network, the creation of a Web Site conti examples of contents useful for students and the development of a Web browser which is the to of this paper.

A preliminary survey on the software tools that are currently used by the visually handica access the Web showed that mastering these tools required a long and difficult training which co be a severe obstacle in our project. The visually handicapped users have to deal with a GUI (graphical layer of Windows), a browser whose interface widely refers to visual meta (Netscape, Internet Explorer) and additional software for translating the contents of the screen i Braille or speech (such as Virgo, Jaws, etc...). This is why was essential to develop a sp browser.

[1] BrailleNet is run by the INSERM-Creare in co-operation with several universities, special schools, companies, and associations.

2 General architecture

Figure 1 shows the general architecture used for the browser ;

- The *document downloader* sends requests to the *Web server* using the HTTP protocol, but can also load HTML files from a hard disk, a diskette or a CD-ROM, or receive HTML pages generated by software engines on the Web.

- a *filter* which processes the source documents and prepares them for presentation in Braille, speech or large print display. Existing Web access solutions for the visually handicapped have clearly shown that a preliminary adaptation of the document can greatly improve its clarity and make browsing through it easier [3,4,5,12].

- Adapted documents are then delivered to the *User Interface* through which the user can read the document and interact with it.

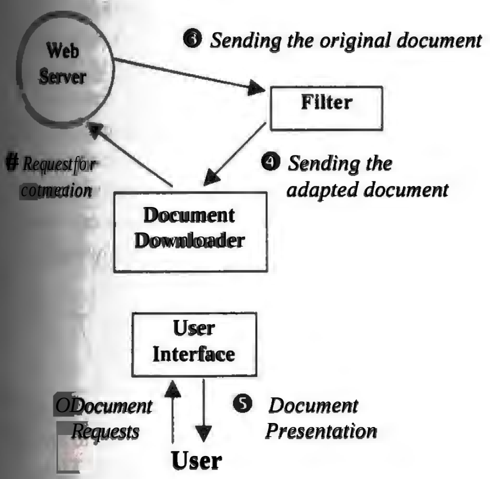


Fig 1 : General Architecture of the browser

The following sections describe the *Filter* and *User Interface*

3 The filtering of HTML documents

3.1 General principles

Internet servers provide information via the unified *HyperText Transfer Protocol* (HTTP), and use the HTML [7] language to describe the structure and content of the information delivered. This information is fairly easy to interpret and transform even before it is presented to the user. Existing access solutions for the visually handicapped have taken advantage of this feature [5,11,12]. There are several levels of transformations can be distinguished :

3.1.1 Simplification

Some elements of little semantic value, such as tags concerning the size, style or colour of text, can be simply removed. Images are also often removed when - unfortunately - no comment is provided to explain them.

3.1.2 Clarification / Rephrasing

Some elements having a very high functional value so that the links must be perceived immediately. Unfortunately neither Braille or speech has a wide range of possible attributes to highlight them. A common solution is to insert a label before the element. For instance, the text browser, LYNX [5] adds a label before hypertext links which makes them clearly perceived on a Braille bar or when they are spoken by a speech synthesiser. In this application, the links can therefore be shown as: `LINK : Design for the Blind`. Links can also be numbered (`LINK 1 : Design for the Blind`, `LINK 2 : Design for the Deaf`).

3.1.3 Restructuring

Even more drastic reformulation is sometimes needed for visually handicapped users to obtain a notion of the global structure of the document and for rapid navigation within it. A summary of links can also be provided at the beginning of the document to facilitate rapid access to them, and the WAB proxy server developed at the University of Zurich also has several more features [12]. Most access products provide a list of the links contained in a document, or a table of contents showing the headings of a document. This table of contents can be preceded by an anchor making it possible to bypass it. This type of adaptation of links implies the complete rebuilding of the document, since new links or anchors must be added. Thus, different transformations can be used single or in combination for each type of HTML element. Optimal solutions generally depend on the user's preference or the display technique used.

3.2 Filtering functions

HTML elements are enclosed in delimiters or tags that clearly indicate the nature of the element. It is therefore easy to trigger a transformation function each time a tag is encountered and to stop it at the end of the element. The process can be applied several times for each transformation. Moreover, HTML is flexible enough to provide a variety of ways to give a semantic content a HTML form. This is the basis of any adaptation based on processing HTML source code. A database can therefore be built that defines the relationships between the various HTML tags and the transformation functions or methods. Table 1 gives some examples of transformations that can be operated on the original sources. These transformation can be operated independently of each other.

Text	• Braille abbreviation • Phonetic value of acronyms for their pronunciation
Headings	• Prefix for Braille display. Ex. : [H1] • Insertion of prosodic markers or voice indicator for speech output
Anchors, Links	• Bracket for Braille display • Prefix for speech output Ex..: "Link" BrailleNet • Numbering the links
Images	• Insertion of an Image indicator. IMG 1

	• Insertion of the caption as a text • Insertion of a anchor allowing the image to be bypassed • If the image has a link associated with it, but no comment insertion of the URL address as a link.
Tables	• Insertion of an Table indicator. Ex. : "Table with links" • Insertion of the caption as a text • Insertion of a anchor allowing the Table to be bypassed
Frames	• Creation of a list of links corresponding to the frames • Prefix for speech output Ex. : "Frame" Menu
Forms	• Insertion of an Form indicator. Ex.: "Form with 3 objects" • Insertion of the caption as a text • Insertion of a anchor allowing the Form to be bypassed

Table 1 : Examples of transformation functions

3 The architecture of the Filter

Figure 2 shows the architecture we have used for the Filter. It uses a *transformation database*, a *customisation interface*, and a *Adaptation* whose function are described below. It processes source HTML documents and outputs adapted HTML documents .

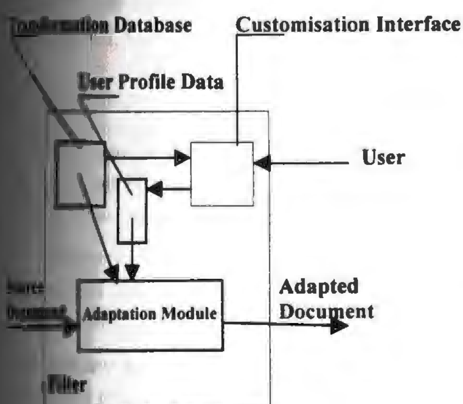


Fig 2 : Filter Architecture

3.3.1 Customising Interface

The document is adapted using the transformations appropriate to each type of HTML ele This is done by a customisation module whose user-interface allows or forbids certain combin of functions, depending on the hardware/ software used.

3.3.2 User Profile Data

The customisation parameters are saved in a data structure and used during filtering.

3.3.3 Adaptation Module

The Adaptation module uses data provided by the database and the customisation module to pro the HTML documents. Each elementary transformation can be broken down in three ste *identification*, in which the HTML tag is identified, *transformation* using one or more adap functions, and *replacement* in which the adapted sequence is inserted into the source docu place of the original one.

4. The User Interface

We first describe the general principles followed in the design of the user interfaces and then examples illustrating how they were applied to build the interface.

4.1. General principles

We have applied a few basic principles in the design of the user interface to insure easy acce reading and navigation functions. These basic principles are generally considered in the desi any computer product, but are particularly important, for an interface system with gre limitations than those of graphical interfaces [2].

4.1.1. Knowledge of the user

Visually handicapped users have been involved in the development to ensure that we took consideration specific knowledge of Braille, the functionality of access products, and to unde certain psycho-cultural aspects of the non-visual approach to reading documents.

4.1.2. Clarity

The content and the nature of the information presented in Braille, speech, or on the screen m easily and immediately perceived.

4.1.3. Coherence

The same interaction methods should be applied throughout the application.

4.1.4. Direct pointing

]i user can to activate a function by designing the item representing this function. On a Braille display, this can be done by clicking on buttons situated in front of the Braille cells. Using a speech based interface, the user can activate the function immediately after having heard its name.

4.1.5 Multimodality

The combination of visual, auditory and tactile communication modes greatly enhances the quality and user friendliness of the interface.

4.1.6 Adaptability

The interface can easily adapt to the specific needs of several users.

4.1.7 User control

The user initiates and controls all actions. For instance, the user can stop the speech synthesiser at any moment and can switch from one modality to another which seems more appropriate.

4.2 The presentations of a document

Computers may be equipped with both a Braille display and a speech synthesiser so that users can read the documents using either the auditory or tactile chaimels. This düal functionality has been implemented, plus the possibility to switch rapidly at any moment from one to another.

Different reading strategies can also be used. A page can be read extensively, skimming over it to understand its structure or to find the links. It is also possible to switch from one strategy to another, tiu Mode Key on the Braille device, and/or on the ASCII keyboard. Thus there are three reading modes (fig 3):

- *Document* Mode: the adapted document is presented
- *Abstract* Mode: the headings and links in the page are shown.
- *Link* Mode: Only the links in the page are shown.

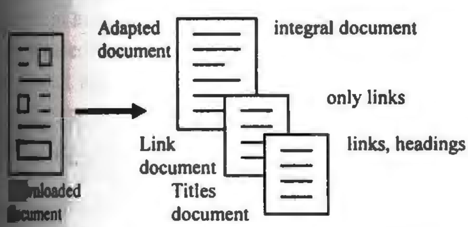


Fig 3: Presentations of a document

4.3 The presentation of Links

Links are essential in HTML and they must be perceived immediately by the user. When data are to be read on a Braille display, the links are presented like blinking which means that they disappear and appear rapidly. The blinking speed can be chosen by the user. A characteristic beep is also emitted when the Braille line is displayed. This solution can be reinforced by filtering, sinc filter can insert brackets just before and after the links (see Table 1)

The user activates a link by clicking on one of the corresponding buttons and the navig functions download the referenced document. If the referenced document is an HTML docu the software invokes the HTML interface. If the reference is an e-mail address, the software invo an e-mail application. If the document is to be downloaded (images, sounds, programs,...) a *Sa As* Dialogue box opens.

Things are obviously very different when reading with a speech synthesiser. Not only a link marked (which is made by HTML filtering, see Table 1), it must also be activated . This means the user must be able to isolate a particular link in a flow of speech produced by reading fin The solution we adopted was for the user to stop the speech synthesiser and focus on the text bei emitted. The user then presses the Enter Key, witch activates the Mode link, presenting the first in the chunk. Thus the user can go through the other links using the Up and Down Arrows. Onc link has been selected the Enter Key is used to activate it.

4.4. Treatment of Multi-frame pages

the frames are transformed into links. Each page referenced by a frame, can be consulted sepa by activating the corresponding link. Thus, the user can go back to the referenced page (Fi Page) at any time, and access another referenced page.

4.5. Screen customisation

The colour, the background and the size of the characters (text, links, controls) can all be chos

4.6. The choice of a language

A Language Key on the Braille device and on the ASCII keyboard makes it possible to switch fr one language to another. Four languages are currently available : English, French, Germa Spanish.

5. Implementation

5.1. Software and hardware environment

The browser has been developed in the environment of Windows 95 and Windows NT using engineering concepts proposed in the ActiveX technology (COM and OLE Automation). programming environment was Visual C++ 4.2 and its extensions (MFC. SDK, APL,...).

The browser has been developed to be fully compatible with most of the Braille devices on French market with 20, 40 or 80 cells, and those equipped with buttons for direct pointing clicking. The speech synthesis system used is the *Proverbe Speech Engine* produced by *Informatique*[2]. It uses a standard sound card.

---

[2] French Society, Toulouse FRANCE

5.1 Using high-level components

Figure 4 shows one way of integrating the application into Microsoft ActiveX technology [6]. It also illustrates how developers tend to design software using sophisticated components whose functions are defined at a very high level. Powerful components can be controlled through rather simple software interfaces in this development environment. For instance, Web Browser Objects can be used to develop an application. It is even possible to co-operate with a complete application such as Microsoft Internet Explorer, since any ActiveX application is provided with a standard software interface based on OLE/COM techniques [9]. Figure 4 shows the software interfaces developed to make the filter communicate with the server and the application, and the HTML [illegible]fication communicate with the user interface.

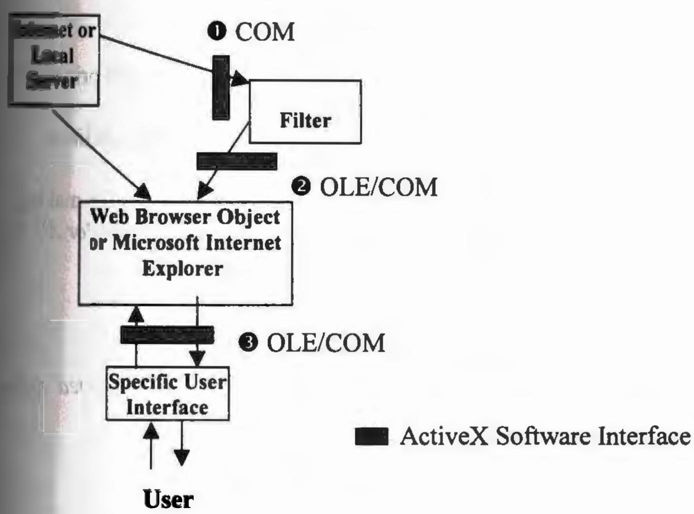


Fig 4: Integration into Microsoft ActiveX

6. Discussion

The principles and methods described here are the basis for developments that have been undertaken at INSERM-Creare as a part of BrailleNet [1]. A beta version of the Web browser has been released which combines visual, auditory and tactile communication modes. As the targets of this project are visually handicapped students from primary schools to universities the interface must be very user-friendly and functional. The proposed approach provides a suitable framework for the development of Internet access products whose main features are full compatibility with current and previous HTML versions, and easy updating as HTML develops by adding new data to the transformation database. This approach clearly separates the adaptation of the web documents from the work on presentation and browsing [8].

7. Conclusion

The applications of the Internet are so numerous and useful that it should be accessible to all users. Although significant progress is being made towards making the Web accessible to people who can barely read a video display, the access still need improvement to make the Internet services as useful to these users as they are to sighted people. Since Web designers cannot be expected to a into account the special needs of very special groups of users, methods have to be devised to adaptation as automatic as possible. This paper describes a potentially useful contribution to problem, which is a fundamental part of R&D in rehabilitation technology [10].

8. Acknowledgements

This R&D project is supported by a grant of the *Fédération des Aveugles et handicapés visuels de France.*

9. References

[1] BrailleNet, *http://www.ccr.jussieu.fr/braillenet/*

[2] Burger D., Improved Access for the visually Handicapped : New Prospects and principals, *IEEE transactions on rehab Engin., 1994,* pp. 111-118

[3] Guide to Writing Accessible HTML: Accessibility for Persons with Visual Disabilities, *http://www.utoronto.ca/atrct/rd/html/htmlvis.html*

[4] Hadjadj D., Agro R. & Burger D., Customising HTML by filtering: techniques for making Internet accessible to the visually handicapped, *3rd ERCIM Workshop User Interface for All.* 1997, pp. 219-224.

[5] Lynx, *http://lynx.browser.org*

[6] Microsoft, *ActiveX Accessibility Conference,* Birmingham-UK, 1996

[7] Schwartre J., *HTML,* Data Becker GMBH & Co KG 1996

[8] Stephanidis C., Access to Graphical User Interfaces by blind people, *Concerted Action Technology and Blindness,* May 1991, p.1-65

[9] Tall E , Ginsburg M, *late Night ActiveX,* Ziff Davis Press, 1996.

[10] The Web and Disabled People, *http://www.w3.org/pub/WWW/Disabilities/*

[11] Unified Web Site Accessibility Guidelines, *http://www.trace.wise.edu/text/quidelns/htmlgide/htmlfull.html*

[12] WAB: W3-Access for Blind And Visually Impaired, *http://inf.ethz.ch/publications/wa.hlml*

Medicine & Gerontechnology

TELE-ASSISTANCE FOR UPPER-LIMB, MYOELECTRIC PROSTHESES

A.Davalli,[1] R.Sacchetti, C.Bonivento,[2] M.Ugolini

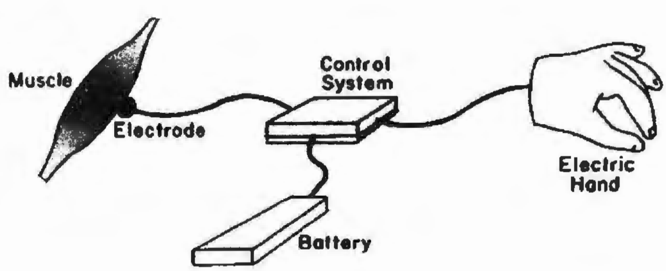
Abstract

The enormous developments made in the last few years in the field of microelectronics and computer programming have permitted the use of technologies in the medical and prosthetic sectors for remote control and tele-assistance. These involve the control and monitoring of a device by means of various communication lines, ranging from common analogous telephone lines to the more recent digital ISDN lines.

In the Research Department of the INAIL Prosthesis Centre in Vigorso di Budrio (Bologna), a project for tele-assistance, applied to the prosthetic field, is being developed which foresees the remote analysis and control of a special type of functional upper-limb prosthesis, equipped with a sophisticated computerised control circuit called MCA (Microprocessor Controlled Arm).

1. INTRODUCTION

The hand is a fundamental organ for demonstrations of creativity that are typical of man. Its loss, either from birth or owing to an accident, is therefore a highly dramatic event for both the practical and the psychological consequences it entails; in such circumstances it is thus advisable to provide a suitable prosthetic device. The extrabody energy functional prostheses currently in use utilize the potential difference that a group of residual muscles is capable of generating by means of an isometric contraction. This electromyographic (EMG) signal is read through skin-surface electrodes fitted in the prosthetic socket, and on this values, the control unit drive one o more actuators (DC motors) that perform the required movements.



[1] INAIL Centro Protesi, Via Rabbuina 16 40054 Vigorso di Budrio (Bologna), Italy

[2] DEIS-LAR, University of Bologna Viale Risorgimento, 2 - Bologna, Italy

As far as the Centro INAIL is concerned, the prosthesis that is generally fitted on pa the Otto Bock myolectrically controlled hand. This prosthetic hand is characterized by a pi movement (that is, to say with only one degree of freedom); it is however well-known for its reliability and good cosmetic which has contributed to making it a standard.

The new generations of upper limbs prostheses developed at the INAIL R&D depa controlled by a microprocessor system that acquired the EMG signals and with a particular co law drive the hand motor. This law is suitable of change to set the best value of its para order to avoid the optimal control of the hand by the patient. The most recent prototype realiz the Centre in collaboration with DEIS- University of Bologna, is the MCA Auto Tuning sys that will be described in this work.

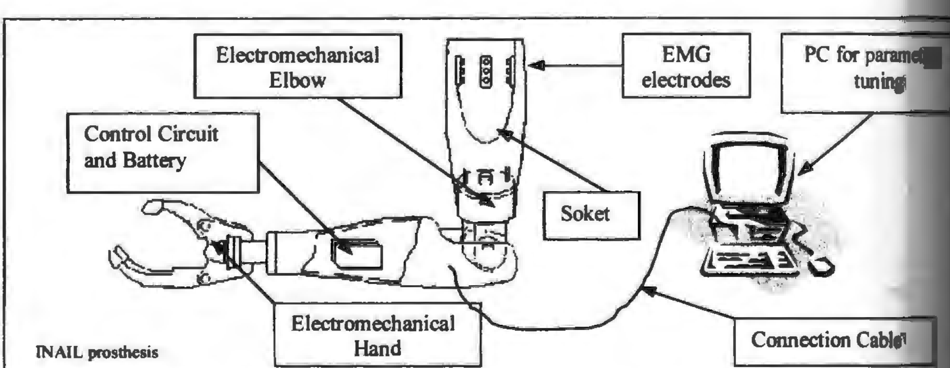
This system is an important step of the global INAIL prosthesis project as it allows en of human skilled operator giving a more freedom to prosthesis users and making easier the wo technicians. In fact, so far the tuning process was realised manually by an expert operator abl understand the intrinsic nature of the prosthesis and specialised in using a fairly difficult pro From now, thanks to a user-friendly graphical interface and to the inclusion of expert-op knowledge, also a low-level user can adjust the parameter values of his own prosth particular, this can help whenever a new setting of the prosthesis is required because of macroscopic behaviour modification. Moreover, also the starting setting, which is gen accomplished by an expert operator and critically depends on EMG signal entity and an level, can be done more easily by using a solution like the MCA Auto Tuning system.

The tele-assistance project foresees carrying out all the necessary operations for the reg of the prosthesis at a distance, by adapting a system of video-communication, implemented with function of "serial port remote control", which permits the data exchange between the PC an prosthesis.

A patient equipped with a MCA prosthesis type, can connect himself telephonically to INAIL Prosthesis Centre and contact a qualified technician capable of regulating the prosthesis diagnosing possible malfunctions, without having to go to the Prosthesis Center personally.

2. INAIL PROSTHESIS CHARACTERISTICS

The considered INAIL myoelectric prosthesis is a new generation multifunction pro that is able to perform hand proportional opening/closing, wrist rotation and el flexion/extension.



The prosthesis is piloted by electromyographic (EMG) signals that are generated owing to residual muscles twitch. Signals are acquired by skin-surface electrodes and are sent to the A\D converter integrated into the microcontroller INTEL 87C196KC. The Analog to Digital converter outputs are then processed to obtain the PWM command of the electrical motor at the desired velocity value.

The control law which is implemented into the microcontoller is a proportional law so that the velocity arm is proportional to the muscle twitch. The control law is customized on each patient using a set of parameters. When the optimum parameters are found, they are stored in a permanent memory and the prosthesis will work with maximum performances.

Using a special software it is possible to change the parameters, for example:

Noise. This value is a measurement of electromagnetic noise that is coupled to the electrodes. The value is subtracted to the A/D converter value so that the disturbance is eliminated.

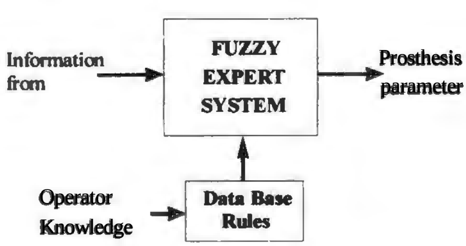
Inactivity threshold. The "I" parameter allows the selection of opening and closing thresholds under which the acquired signals are not processed. In practice it is necessary that at least one of the two signals is greater than the "I" value to move the prosthesis. The parameter has been introduced to filter out spurious signals that are generated when the patient moves the body without wanting open or close the prosthesis. Without this threshold the spurious signals would move the artificial arm. If the parameter value is too low then also the "physiology noise" produces an imperceptible movement that augments the power consumption.

Maximum threshold (M). The M value assigns the upper power limit over which the motor gives the maximum power value. The threshold allows the patient to reach the maximum velocity prosthesis even if the signal is weak.

Extensor gain (E) and Flexor gain (F). These parameters assign the gain connected to extensor and flexor signals respectively. These values are used to level by software possible difference between the two signals.

3 THE TUNING SOFTWARE PROJECT

A particular software package allows automatic or manual tuning of myoelectric prostheses by using a Fuzzy Logic Based Expert System. Such a system first acquires information regarding the actual artificial arm working and its desired behaviour, then applies a specific set of rules representing expert operator know-how to define the parameter set satisfying the patient requirements.

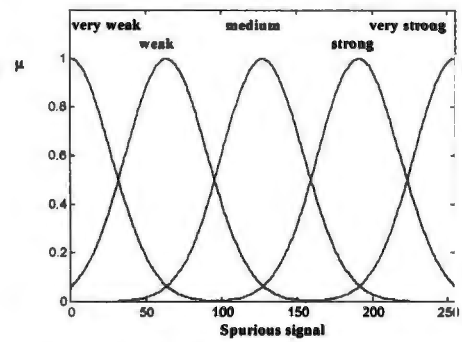


General scheme of a Fuzzy Logic Based Expert System

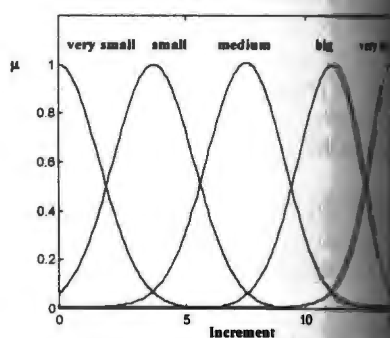
A general scheme of a Fuzzy Logic Based Expert System is shown in the previous picture. The software package has been developed using Visual Basic programming language for Wi 95 operating system. The final product is so simple to use that also the patient himself can i without being helped from an expert operator.

The more innovative part is the one relating to a Fuzzy Logic structure which implem the Expert System. This expert system is called by two controllers dedicated to calculate Ina threshold and E and F gains respectively. In the following we look for example the Inac threshold Fuzzy Logic Controller.

The input of Inactivity threshold Fuzzy Logic Controller is a value resulting from spu signal acquisition process while output is the increment to be assigned to the parameter value. input space has been represented by five Gaussian fuzzy sets uniformly distributed on the range and called "very weak", "weak", "medium", "strong" and "very strong" respectively. fuzzy sets represent the membership functions degree (μ), as represented into the diagram.



Membership function degree to represent spurious signal (input variable)



Membership function degree to represent Ind increment (output variable)

Similarly, the output variable is a gaussian fuzzy sets which are e labelled as " very sm "small", "medium", "big" and "very big" respectively.

The implemented rule set is represented as follows:

1. If spurious signal is very weak then increment is very small
2. If spurious signal is weak then increment is small
3. If spurious signal is medium then increment is medium
4. If spurious signal is strong then increment is big
5. If spurious signal is very strong then increment is very big.

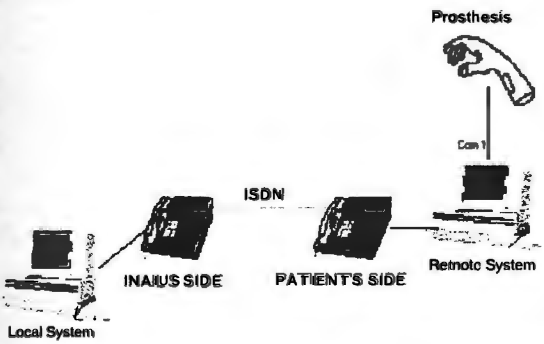
Parameters determining the macroscopic behaviour of the prosthesis are related one an In order to simplify the project we decided to calculate one parameter at a time but followi well-defined order to not neglect the mentioned interdependence.

4 THE TELE-ASSISTANCE PROJECT

MCA Auto Tuning Software has been integrated with the necessary modules to develop a telematic service system for the disabled. In this way the control software may be used for tuning both when the prosthesis is connected directly to the PC and when the prosthesis and PC are physically located in different places. This solution makes it considerably easier for those disabled persons who would otherwise have to face difficult, costly travel to specialised centers even for simple check-ups.

Communication via ISDN (integrated Service Digital Network) has been chosen for this purpose, which allows several services to be combined including audio, video and data in a single transmission band, at a speed of 128 Kbytes/sec.

The telematic service system planned is illustrated in the figure below, showing the local and remote systems - of the operator and prosthesis wearer, espectively - communicating over the network.



General diagram of the telematic service project

Various possibilities have been evaluated for solving the telematic service problem, each with its own positive and negative aspects.

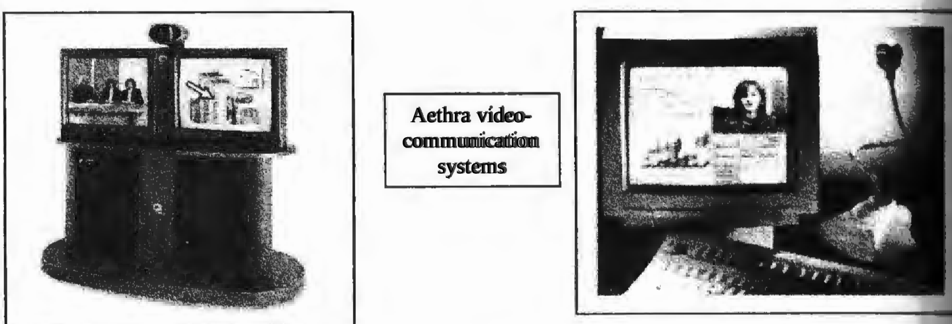
One of the possible solutions implemented, which satisfies system specifications, is to use Application Sharing: in this case, the MCA Auto Tuning software and Aethra video communication support need simply to be distributed among the various remote locations; the expert operator can use the application from the headquarters as though the patient were on site. Remote application sharing is made possible by various programs available on the market, including Microsoft NetMeeting, supplied directly by Microsoft with the latest versions of Windows 95 and Explorer.

The other solution implemented is conceptually different from the above, since it avoids distributing the control software. This is instead centralised, and uses virtual communication between the PC and prosthesis. This requires establishing remote serial communieation via the ISDN network. This new solution streamlines data traffic between the two systems - quite heavy in the case of Application Sharing, due to the ongoing delivery of video maps from the remote to the local system. In this case only the prosthesis parameters are transmitted, making it faster and more effective, it also avoids the security problem inherent in distributing software, which in this instance is used and handled only by expert personnel.

This type of system has been possible to develop thanks to tools by Aethra Telecomunicazioni, a leading European firm in the video communications field, which supplies a

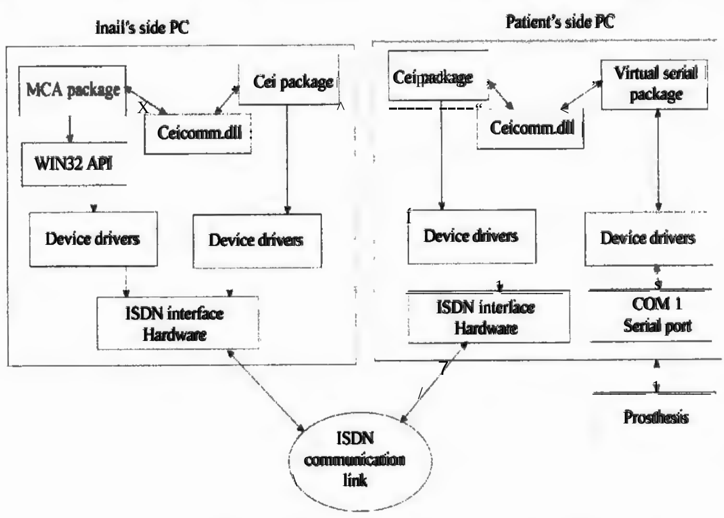
set of hardware equipment to interface with the ISDN network and software through the
(Software Development Kit) libraries to transmit serial data.

The MCA Auto Tuning Software processes the prosthesis parameters that are exch
over the serial connection. If the patient is physically present in front of the operator, the
connection is to the physical port of the PC on which the technician is working, in the case
remote patient, the operator works on a PC with a "virtual serial" port mapped by special sof
on the ISDN line. One especially significant factor in developing the project was the crea



drivers to manage the serial communication between the various remote terminals and the IS
network, Two software drivers handle this data: the Portdrv.vxd driver required on both the
and remote sides, and the Caronte.vxd driver required only on the remote side. Portdrv.vxd i
an additional serial port on the system, COM6, which is not physically present on the PC but
behaves for all intents and purposes like a real serial port. The programmer simply uses this vi
serial port to communicate between the PC and the ISDN network. The Caronte.vxd (d
manages instead the interface between the physical serial port connected to the prosthesis and
ISDN network.

The following diagram shows the level and manner in which the various Aethra component
in terms of routing audio/video calls and managing the "virtual" serial communication bet
prosthesis and PC in the two locations. The two drivers work on the first software level, thus
than the user level, where the video communication interface is managed.



Data and Control flow diagram through components

5. RESULTS

The telematic service project thus developed has been tested on a few patients with highly satisfactory results, and has made it possible to verify the effectiveness of the software both in local connection with the prosthesis and through the video communication system. In the case of a remote connection, one of the greatest hurdles that have been overcome had to do with maintaining efficient communication between the prosthesis and the control software.
Indeed, the communication requires that information pass through various types of software and hardware protocols, leading to synchrony problems due mainly to delays introduced by the components. Imperfect synchrony could cause serious difficulty, both in terms of video images of the prosthesis not matching the signals reaching the PC and, in more serious cases, the loss of information along the line.

6. CONCLUSIONS and DEVELOPEMENTS

The presented system, is a useful tool both for operator and amputated themselves. Amputated itself have in fact a more freedom degree since he can auto tune his prosthesis without going every time to the Prosthesis Centre; expert operators can manually tune a prosthesis by using a user friendly and easier program and using a PC video-communication system we can see and talk with the patient. Therefore using MCA Auto Tuning package the present troubles and outlays for the patient can be dramatically lowered since any artificial arm resetting requires only a few easy adjustment staying at home.

The adoption of this program on a large scale may yield considerable economic benefits and improve the service quality supplied to the prosthesis user. In fact the time required to set the prosthesis parameters is remarkably reduced and, consequently, working time of technicians is reduced too, decreasing costs of prostheses producers and providers. Moreover, the software can be distributed to all users using tele-assistance service and network service decreasing mobility and assistance costs and providing a better service.

Telematic service, limited in this case to controlling myoelectric prostheses, may easily be extended to service wheelchairs for the disabled, where once again a serial communication is needed to set and control the parameters that ensure smooth operation.

References:

[1] A.Davalli, MCA3 Ver.3.1 - Microprocessor Controlled Arm - User manual.
INAIL Centro Protesi Ricerca e Documentazione, 1996

[2] D.Brubaker, The Huntington Group, Fuzzy-logic basics: intuitive rules replace complex math. EDN, June 1992

[3] F.Mascaretti, C.Panini, Software Developement Kit SDV8000 - User manual.
Aethra Telecomunications, 1995

[4] S.Terenzi, Automatic Tuning of Microprocessor Controlled Arm – Degree thesis.
University of Bologna, 1996

IT SUPPORT FOR DIABETES

Wolfgang Reh, Christoph Cronrath, Helmut Wenzel, Karl We

Boehringer Mannheim GmbH, Sandhoferstr. 116,
D-68298 Mannheim, Germany,
Fax +49 / 621 - 759 3101

Abstract

Diabetes is a widespread chronic disease with even growing importance in the future. The ma ment of diabetes, which means prevention and/or treatment, can be improved and mo efficient by employing the tools of information technology (IT). The authors describe the IT requirements of the individual tasks of diabetes management and describe tools and solutions for meeting those requirements.

1 Introduction

Diabetes mellitus (DM) is a complex, chronic metabolic disease characterised by elevated glucose levels. DM essentially takes two forms: IDDM and NIDDM. Insulin-dependent mellitus (IDDM) usually begins in childhood, adolescence or early adulthood. The insulin ou the pancreas is no longer sufficient or may stop altogether, and blood glucose rises to threatening levels. Non insulin dependent diabetes (NIDDM) - also labeled Type 2 - occu in middle and older age groups. Often associated with overweight (obesity) and reduced activity, diabetes is the result of insensitivity of the body's cells to insulin (insulin resistance).

Diabetes mellitus is a common disease: more than 135 million people may be affected world According to projections of the WHO, this number will rise to 220 million in the year 20 possibly will reach 300 million prevalent cases by the year 2025 [1]. DM is also a costly that imposes great burden on the patients, their families, the public health, the society plies, mainly due to the appearance and the progression of its chronic complications. Based on the expected number of people with Type 1 DM in Germany today, health insurance has to pay about 170 billion Deutsche Mark approximately to treat DM and its complications in this cohort. In this regard, diabetic patients represent 8% of ambulatory care of patients and consume 20% of the health expenditures in ambulatory care. Improving quality of care with the consequent decrease of chronic complications now (only in the treatment of Type I patients) could help health insurance to save about 2.1 billion DM per year [2]. In 1995, 15% of the US health care budget [3] was spent for the treatment of diabetes.

Improved glycaemic control substantially reduces both the development and the progression of diabetic microvascular complications, i.e. lowering blood sugar reduces relative risk of retinopathy (76%), nephropathy (50%), neuropathy (60%), and cardiovascular disease (35%) in people with Type 1 DM [4]. Thus, measures that improve quality of care and glycaemic control might simultaneously improve the quality of life of people with DM and decrease the short- and the long-term cost of the disease.

Successful quality assurance approaches need specific prerequisites:

- a clear definition of management goals (on the basis of medical parameters such as bG, HbA1c, blood pressure, etc.) that should be reached,
- medical tools and know-how for prevention and/or treatment, and
- technical tools to satisfy the needs of permanent feedback loops that are needed to close the gap between planned goals and actual readings.

As this control takes place at different levels of the disease management process (patient level, hospital, GP, health care system, health insurance), different degrees of expertise and qualification have to be taken into consideration. Based on a growing body of knowledge and experience in the treatment of diabetes, treatment **standards** and disease management **goals were established.**

Tight metabolic control of diabetics means that a flood of data must be handled. This data management can be improved by **computer** and **telecommunications** support. Computerisation makes storage, evaluation and transfer of data effective, thus contributing to a more efficient disease management by a faster feedback of results.

2 Tasks of diabetes management

2.1 The diabetes team

Diabetes is a chronic disease that requires constant medical supervision and training of the patien acute metabolic disturbances and late complications are to be avoided. Treatment of diabetes is responsibility not only of doctors in various specialities, but also of nurses, dietary advi psychologists, etc. After the initial diagnosis the patient is often hospitalised for insulin stabili and training in a diabetes ward. The patient then visits the diabetologist every three months.

IT solutions can be applied to:

- helping the patient to meet the currently optimal blood glucose target,
- training the patient,
- adapting the treatment strategy,
- avoiding late complications, and
- optimising documentation and communication.

2.2 Self-management by the patient

After appropriate training, patients should be able to make their own decisions about insulin do meals and physical activity in accordance with their current state. Patients are confronted wi complex task of maintaining continual blood glucose regulation on the basis of a few (3-7) d blood glucose measurements and their physical condition using a treatment strategy agreed their doctor.

Self-management of diabetes was made possible by the availability of hand-held mini instruments with these components:

- puncture aid to obtain the necessary amount of blood,
- blood glucose determination system (test strips with or without instrument),
- insulin injection aid or insulin pump,
- tools for making short-term decisions (treatment plan),
- diabetes logbook for documenting metabolic status including information on meals, exercise and events, and
- medium for communicating with diabetologists.

To make the system easier to operate and to improve the quality of the evaluation, manufacturers have developed instruments that measure blood glucose. By using micromechanics and integrating the electronic components (CPU, storage elements, interface components, etc.) into a single ASIC, small battery-powered blood glucose measurement instruments have been developed that are portable and robust. Instruments with "measure," "record," "evaluate," and "communicate" functions are now available and empower the self-management by the patient.

2.3 Training the patient

Properly trained patients have to make their own decisions about their treatment strategy. To apply this treatment successfully, however, the patient and the care team need more application-related knowledge about the effect of the various influences on the result.

These demands can only be met by modern training methods. Especially important here are interactive education and training programmes that not only employ the written or spoken word, but also allow the user to steer the learning process in accordance with the user's own experience and needs.

By simulating the individual metabolic situations, for example, the effects of insulin, meals, and exercise can be made clear in a very short time to the patient. Ideally, simulation models allow the user to deal with the real patient data in order to forecast effects on insulin regimen, meal planning and physical activity.

2.4 Diabetes management and treatment optimisation

Self-monitoring of blood glucose by the patient produces a mass of data. This complex mass of data must be organised and structured and undergo an analysis appropriate to the situation.

Dangerous trends and problems in the circadian rhythm have to be recognised at appropriate times, and data suitable for decision-making must be processed. Together with the patient, doctors and diabetes advisors draw up individual dose adjustment strategies, and they need quick feedback on the results of treatment modifications. At consultations between doctor and patient, documentation aids and tools for making the evaluation procedure easier to understand are needed. Graphic presentation of data with subsequent printout of reports will become indispensable.

2.5 Process optimisation

For diabetes professionals the question of quality assurance and quality and cost managemeili becoming increasingly relevant. Quality simply must be measured, documented and mainta demonstrably favourable cost. To obtain quality data that can be compared with others, sta such as the Diabcare Basic Information Sheet (St. Vincent) should be applied as far as pos Using data processing tools, result and process quality indicators can then be derived from documented diabetes-relevant quality data. Analysis of the efficacy of treatment methodi a optimisation of standards and so reduces the cost of diabetes treatment.

2.6 Diabetes in telemedicine

Telemedicine is the use of electronic communication and information technologies to provide support clinical care from a distance. Diabetes is an excellent example for the applicatmi telemedicine. Everyone involved in diabetes treatment must work very closely togethw. advising the patient must have access to current information about the patient. When adjusti treatment, frequent consultations with the treating health care professional are necessary! Re data transfer with advice given by telephone can avoid serious complications to the patient and reduce treatment costs.

3 Integrated computer-aided solutions and services

For the diabetes management tasks outlined above, integrated computer-aided solutraBj necessary. To be effective and efficient, the integration must include all essential stages of d management. The matched solutions and services described below are an integrated contribuliiil the pharmaceutical and diagnostics manufacturer Boehringer Mannheim to diabetes managemtll

The range comprises:

- Accutrend® DM: BG measuring instrument with built-in data management and PC interface
- Camit® for Windows™: data management and analysis tool for self-monitoring data
- Diabcare® for Windows™: software for result monitoring and diabetes documentation
- Interactive education and training programme

- Accusim®: simulation of outcome and economic effects
- WellMate™: telematic diabetes management

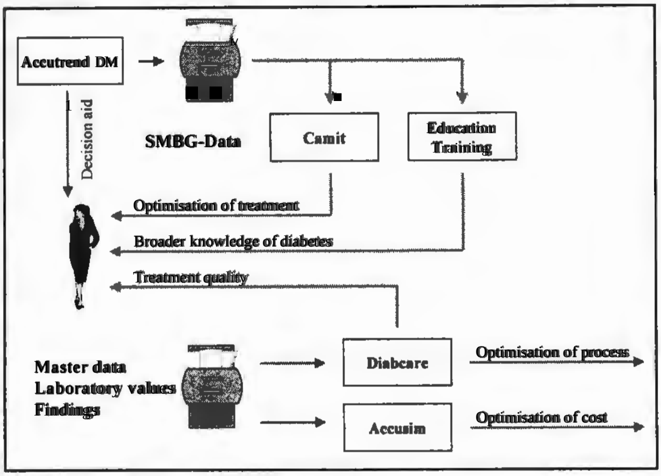


Fig. 1: Integrated diabetes management systems

3.1 Electronic logbook

Accutrend® DM incorporates the functions of a blood glucose measurement instrument and an electronic logbook with graphic and tabular functions. Characteristics are:

- fast BG measurement with small instrument,
- user-friendly interface with plain text messages and data entry via number and function keys,
- manages and stores the metabolic information including meals, insulin, exercise and events,
- large display for graphic presentations, tables, statistics as a decision aid, and
- indicates trends and gives warnings in critical situations (hypoglycaemia).

Fig. 2: Meter and Elect. Logbook; Accutrend® DM

3.2 Analysis tool for self-monitoring data

The analytical tool Camit® processes the measurements and logbook information stored in measuring instruments such as Accutrend® DM. These data can be uploaded into the Ca database through an interface cable supplied by the manufacturer. Camit helps doctors and pat to construct a database of measurements and events with practically no time limits.

Even with just four blood glucose measurements per day, there are 360 measurements per qu Condensing this information into meaningful graphs and documents entails an enormous amou work. This is where Camit® can help decisively, improving communication between docto patient and motivating the diabetic to make the personal effort needed for self-monitoring.

A standard procedure for discussing the Camit® analysis can be agreed upon by the docto diabetes advisor and the patient. The evaluations can be organised in user-definable reports.

Camit® offers analysis functions such as trend graph, standard day, standard week and adapt Event-specific analyses lead to faster improvement of insulin adaptation schemes. Tools recognition of trends are part of the package. Statistical analysis functions assist assess adjustment to the selectable target BG range.

3.3 Documentation and outcome monitoring

The PC software Diabcare® is an important element in implementing the St. Vincent Declarati diabetes institutions. Diabcare® provides a standardised basis for long-term strategic monitori results and adjustment quality. The aim is to improve diabetes treatment and avoid complications. The Diabcare® data set is based on the Basic Information Sheet and is arra the following main groups:

- laboratory data including lipid values, blood pressure, HbA1c and BMI (calculated),
- treatment data (including insulin dosage scheme),
- complications (eyes, kidneys, feet, neuropathy),
- training data,

* data on quality of life, and
* patient master data.

Outpatient data are arranged by visit date, with the most recent values available displayed on screen. Thus Camit® and Diabcare® permit complete evaluation and documentation of all diabetes-related patient information.

3.4 Interactive education and training programme

To promote interactive knowledge transfer, Boehringer Mannheim and the Diabetes Research Institute in Karlsburg have begun a collaborative project aimed at combining the Karlsburg Diabetes Management System and the Accutrend® DM and Camit® systems into an interactive education and training programme for diabetics.

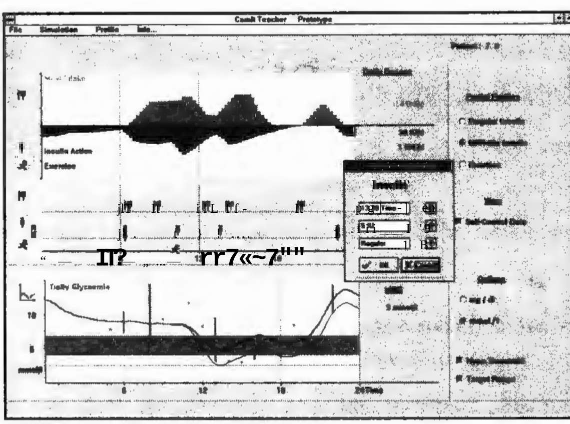


Fig. 3: Interactive education and training programme

Using this programme, individual patients will, for the first time, be able to use their self-monitoring data to simulate the effects of changes in therapeutically important parameters on their 24-hour blood glucose profile by means of differential equations. The patient will enter the dose, the type and times of the insulin injections, the quantity and composition of the diet, and the physical activity into the programme in any desired sequence.

3.5 Telematic diabetes management

In order to further improve the communication between health care professionals and patients, especially in situations like pregnancy, newly diagnosed diabetics, etc., Boehringer Mannheim and

Nokia are working on a new interactive telematic diabetes management concept. It combines latest telecommunications technology with state-of-the-art diabetes disease knowledge. WellN introduces a new diabetes-specific software for a cellular phone that enables a diab conveniently track diabetes-related data such as blood glucose values, insulin dosages, etc. patient simply presses a button to transmit data to a database where the information is conti collected and automatically analysed. This database, of course, will meet the highest stand data security and confidentiality. Authorised physicians will be able to request key inform the Internet using standardised graphics and statistical functions to support decision mak Diabetics can get their individual data anywhere and at any time using a mobile phone. It is possible to retrieve more detailed analysis and education programmes by Internet or fax.

This concept is designed to support health care professionals in their daily work and to diabetics to manage their disease by themselves. It is an innovative communication and management tool linking physicians and patients; however, it cannot replace regular physi consultation.

Currently, the first clinical and technical tests are being conducted in Finland.

3.6 Simulation of clinical and economic outcomes

What decision makers need is an extrapolation of health and cost outcomes of their popula assuming they will introduce new therapies, and here increasingly they expect a contribution industry. The payer also needs to find out where the resources may be used up in an inefficien Maybe there are no additional resources needed when improving quality of care, and they must reallocated according to perceived risk of different population strata and expected benefit. To d decision makers need to understand how a given population – which can be described by gender, risk factors, and ethnic group – will behave over time. This is a very challenging because uncertainty of events and a highly complex and nested process has to be taken into acco

The likelihood of reaching the goal can be described by disease model scenarios. These sce are based on the current characteristics of the population, the financial characteristics, and business situation. Introducing a disease management concept or a quality assurance pro

Édd change the situation in a way that the planned goal can be reached. Measuring intermediate nobles such as blood pressure, metabolic parameters (HbA1c) etc., will prove the effectiveness of the quality assurance and lead to the confirmation or the refocusing of the intervention measures, hil it cannot show the impact on the final outcome in terms of expected complication rates, life apectancy and, last not least, financial impact to the business.

The Accusim* 2.0 Diabetes Disease Management Model is a real-time simulator of the clinical ntcomes and economic effects of the long-term complications of diabetes under different diabetes Mgement and screening strategies. The model analyses the development of seven important amplications of diabetes: nephropathy, retinopathy, acute myocardial infarction, stroke, putation, hypoglycaemia, and ketoacidosis. Life expectancy, incidence and prevalence of amplications, and lifetime costs of diabetes management and complication treatment are calculated for specified cohorts. Results are presented as summary results, as well as divided by submodel, llowing identification of the complications driving the costs of diabetes care.

Disease models thus enable a higher degree of complexity to be modelled than is encountered in real-life populations.

4 Concluding Remarks

Diabetes is a chronic disease of utmost importance to public health. Because of its complex metabolic problems and extended data management needs, IT solutions become indispensable. These IT solutions must comprise all stages of diabetes management and have to support patients and professionals in an integrated manner. Telematic diabetes management and disease models simulating clinical outcomes and economic effects now become available.

5 References

(1) Gruber E. et al. (eds.): The Economics of Diabetes and Diabetes Care. A Report of the Diabetes Health Economics Study Group, 1997

[2] Results of a simulation based on the number of type II cases in Germany. The treatment cost of standard therapy and improved therapy are outcomes of the diabetes disease model "AccuSim". A calculation of a German Health Insurance Company comes up with similar data.

[3] ADA, diabetes facts: http://www.diabetes.org/adami/beta/information/diabfact.htm

[4] Results of the DCCT, see New England Journal of Medicine, 329 (14), September 30, 1993

SMART HOUSE FEATURES IN THE HOMES OF ELD-ERLY - An End User Oriented Approach

Mervi Lehto

Abstract

The answers of attitude towards new technology are split into groups in order to recognise end user needs. Two thirds of those who answered reported willingness to learn to use new tech ogy, if needed. One tenth of them is already acquainted with new technology. A fifth of them not want to have it at home, if possible. Safety is most important among elderly, as well as all aged people. The same amount of elderly people - it is roughly one third - want have equipped home office as the people of other age groups have. Elderly suppose to be served formation thanks to the new technology.

1 INTRODUCTION

New technology of elderly is expected to aid in solving the problems of growing group of citizens in need of assistance. Many solutions for smart housing and independent living, and technology, have been created to fulfil these expectations [1]. A questionnaire study was carried to make a draft of the needs of elderly for housing and technical apparatus from the view building developers. Willingness to have and use new technology at home including cost and attitudes was surveyed.

This paper is concentrating on the prior findings of the end users' opinion on the smart house tures related to the means of independent living and gerontechnology. The results will be disc in more detail in the 1st International Congress of CIB Working Commission W098: "Intel (IB) and Responsive Buildings" [8]. The study gives also an overall view of the future style ágto be discussed in some other connection. At least, it is listing spaces, services (valued added services included) and domestic appliance senior citizens are favouring.

Utilisation of ICT (Information and Communication Technology) has been a practice in many cases already for years in business and everyday life, in Finland as well as in other information intensive small nation. That is why, in Finland data is available on the use of ICT based on the operative experience. In some cases, the technology was either so new or so rarely used, that those who answered were lacking experiences of it. Those who answered were at the age of 55 to 70 years. Accordingly, the study had a futurist approach. The aim was on the needs, which can be employed in coming 5-7 years.

It is assumed that the reader is familiar with the definitions of the IT, ICT or telematics. The cases of independent living and gerontechology can be found in wide literature of the subjects (see [2], [3], [5], [6], [12]). There is a consensus of making as small difference between healthy and other people as possible, and following this principle in technical solutions accordingly. A modified piece of domestic appliance can work as an auxiliary mean for elderly. Compared to tailored made technical aid this solution would reduce the cost of products, as well as improve social equality.

IB meets the spatial demand of the modem society with help of IB concept by means of flexibility, activity, integration, information, interaction and designs based on correct understanding of natural laws. Although, the intelligent working environment can be created by the means of organisation, architecture and building technology, for the time being, ICT is one of the most powerful tools for it. The current definitions of IBs (see [4], [9], [12]) emphasise the meaning of user, when the qualities of IBs are defined, and the other means of implementing IB features than the ICT, such as active features, indoor design, etc. not forgetting the electrical brains of the building.

1 METHOD

The questionnaire covers such subjects as follows: need of services, domestic appliance and electronics, spaces and their quality [8], ICT, home and building automation, smart house features, health care equipment. The formula was sent to a randomly selected group of tenants, in the central part of Finland (Pirkanmaa), where are located an industrial town and agricultural communities. The share of returned formulas was 32 %, it is, over 300. The Finns born in 1925 - 1939 are those, who rebuilt the country after World War II, and had still been of working age during the good economic trend in 1960's and 1970's. The answerers' educational status is not very high (57 % of them had only the basic schooling, rest of them had studied some more). The average schooling status of the whole age class in Finalnd is some higher. The share of female was 54 % corresponding that of the average. The portion of those still working is 17 %. The share of 44 % is pensioned because of their age and that of 18 % because of illness. Rest are unemployed, working part time, etc. One third is reporting good health. Elderly live either with spouse (58 %) or single (37 %).

Much expert advice was involved in the formulation of the questions (of 100 variables), and the principles of quality management (ISO 9000) were followed, ending up to a 12 A4 pages long formula. They were divided into categories according to age (55-59, 60-64 and 65-70 years), gender and working status: still working, lately pensioned and already some years on pension. A questionnaire is a good method to make an overall draft of the factors of complex system. Only, it has been criticised to be unreliable due to mistakes in the formulation of the questions and the lack of knowledge among those, who answer. Many facts give proof to the end users' correct and rational opinions.

The conclusions are drawn by a comparison of [a] users' opinions of various groups mentioned above, and of the results of this and comparative studies [b] in Finland and [c] abroad. In the first place, the comparison was made to the results of a similar questionnaire carried out in the City of Varkaus (in 1991) among people in age of 45 to 64 years [7]. Secondly, the results of a questionnaire in the City of Espoo (in 1990) among all aged people (over 700 answers) followed [11]. Furthermore, the TIDE projects' results [2], [3] have been taken account.

3 CONCLUDING RESULTS

The study proves a positive attitude towards the use of new technology (Table 1), which is in agreement of the other studies mentioned above. In general, Finns are favouring new technology, but the results of comparative TIDE projects [2], [3] confirm the finding to be international. The group of oldest, the one with poor earnings or education had had less experience on use of IT, etc. They also decide, more often than others, on alternatives expressing hesitation of usability of technology. Most of those who answered (70 %) want to learn to use new technology, if necessary. One tenth (9 %) is already acquainted with it. The share of those who do not want to learn it is 16 %

Over a half (54 %) expect the municipal or governmental authorities to pay the extra costs, which their ageing is causing, such as technical aid. A quarter (27 %) is ready to pay them themselves. The share of 15 % is willing to trust on insurance.

Domestic appliance similar to other families is a standard in homes of elderly. The only exception of the rule is a dishwasher, which is on their shopping list. Only the share of 31 % of elderly had it, while that of younger families is 71 % [11]. Elderly want to have also such luxury as spa, integrated vacuum cleaner, etc. New technology is supposed to help in everyday family routines, especially a well equipped kitchen (Table 2.). That is the case with safety facilities, too. The kitchen becomes more important with ageing. Mechanical air conditioning, water leakage controls and lighting control are popular. A natural light controlled lighting becomes important with age, while dimming (32 %) is important for other age groups [11]. Most (60 %) like to keep home control as it is. Other forms of home control (popularity in parentheses) are: remote control units (21 %), a control panel (13 %), door side panels (8 %), speech control (7 %) and a computer (3 %). Every fifth of the oldest age group plan to make changes in their house or apartment. Still, their need for new technology is not good, and for example movable wall is not in favour (66 % say never needed) among them. Others need movable walls mainly in 5 to 10 years' periods (50 %) [11].

Every fourth of those who answered has a home office with ICT facilities and the average size of 14 m². The number of home office could be bigger, while every third would like to have a one. For comparison reasons can be mentioned that every second Finnish well educated office worker has a home office [10]. When a home computer (14 % having, 12 % more wish to have) is not any more needed for working, it will be used for such services as health monitoring, internet service access (like banking, ticket booking) and home data bases. People just pensioned do not consider safety and health care equipment as necessary as elders. In fact, they are feeling better than those still working. Actually they seek for service and activities.

4 CONCLUSION

In general, it can be stated based on the results that a big share of the elderly is willing to have and use new technology at home, which encourages the contractors to build smart housing for elderly.

Table 1. The willingness to use technology for daily household tasks.

Alternative\a share of answers (%)	All	age of 55-59	age of 60-64	age of 65-70	Female	Male
Yes	28	29	34	21	27	29
No	19	19	17	18	20	17
Depends on solutions' quality	38	43	40	33	34	[illegible]
Depends on other help available	16	8	9	28	18	[illegible]

Table 2. The share of those who are willing to have smart home features or health care facilities at home.

Active Structures and Intelligent Building Services	%
intelligent kitchen stove (temperature setting and cooking timing)	24
movable kitchen cabinet with push buttons (moves up and down)	17
remote control of sauna stove (phone or computer control)	13
kitchen cabinet with adjustable worktop levels and push buttons	10
motorised windows (remote controlled opening and shutting)	8
motorised doors	8
motorised gates	6
motorised cellar (filling up and taking off with push buttons)	6
remote controlled curtains (on/off)	5
other	3

Home Health Care	[illegible]
safety kitchen stove	[illegible]
safety medicin reminder	[illegible]
movement monitoring	[illegible]
washing wc-seat	7
mattress monitoring	[illegible]
healthy wc-seat	4
video phone, video control	[illegible]
other	[illegible]
nothing	[illegible]

5 REFERENCES

[1] ANON. A. 1996. Handynet database (CD-ROM). Handycom. Brussels. (email:imex@dx-net.fr)

[2] ANON. B. 1995. Concept of Automation and Services for People with Special Needs. User Requirement Defi[illegible] tion. TIDE Project 1068. CASA. 169 p.

[3] ANON. C. 1994. Telematics for the Integration of Disabled and Elderly people (TIDE). Background to the Wo[illegible] plan. Based on the work of the technical panels and a wider consultation with sector actors. Brussels 1995. 150 p.

[4] BEDÓS J. M., PUIG G. A. & FONTANALS R. LL. J. 1993. Intelligent Buildings and Areas. Definition [illegible] emerging concept. Institut Cerdá. Project INFRA. 32 p.

[5] COLOMBI D., RAFFERTY J., STEYAERT J. 1993. Human Services and Information Technology: A Eu[illegible] Perspective. ENITH European Network for Information technology in the Human Services. 208 p.

[6] GLASTONBURY B. 1996. Editor's Introduction: Dreaming with Care. Dreams and Realities. Information Techn[illegible] ogy in the Human Services. Papers from the HUSITA 4 Conference held in Finland 1996. Stakes. National Res[illegible] and Development Centre for Welfare and Health. Helsinki 1996. 216 p.

[7] KÄRNÄ, S. 1992. Keski-ikäisten tulevaisuuden näkemykset ja toiveet omasta vanhuuden palveluista 45-[illegible] vuotiaisiin varkautelaisiin kohdistunut kyselytutkimus. Kuopion yliopiston ja Varkauden kaupungin [illegible] hteistyöprojekti. Varkauden kaupunki. 90 p. (in Finnish)

[8] LEHTO M. Smart House Features of the Homes of Elderly. Intelligent and Responsive Buildings. CIB Wo[illegible] Commission W098. 1st International Congress. March 29th - 30th 1999. Brugge, Belgium. (Abstract)

[9] LEHTO M. The 'Intelligent Office' Concept Makes the Difference. Nordic Architectural Studies Review. No[illegible] Arkitekturforskning 1996:1. pp. 61- 72.

[10] LEHTO M., FLINK S., KARJALAINEN S., KOVANEN K., PULAKKA S. 1997. The Feasibility of the Int[illegible] Building Features in Office Buildings. The User Oriented Quality, the Indoor Air and the Cost Effective[illegible] Building Techn. Espoo. April 1997. 256 p. (in Finnish)

[11] LEHTO M., JANTUNEN J. 1991. Älykäs rakennus. Markkinatutkimus. Otto Wuorio Oy. VTT/YRT. Luotta[illegible] linen tilaustutkimus. 1.2.1991. 16 s. + liitteet 56 s.

[12] RANTANEN J. Human-technology interface: Problems and challenges. HUSITA 4 Conference, Information T[illegible] nology in the Human Services: Dreams and Realities. Rovaniemi. 11-14.6.1996.

[13] YASUYOSHI M. Advanced Intelligent Building in Japan. FUTURE/BUILD '93: "Building for Comp[illegible] vantage". IBI. Chicago Illinois. Oct. 25-27.1993. 11 p.

Session V

Access to Hypermedia

f

THE USE OF NON-SPEECH SOUNDS IN A HYPERMEDIA INTERFACE FOR BLIND USERS

Sarah Morley[1], Helen Petrie[1], Anne-Marie O'Neill[1] and Peter McNally[2]

The use of non-speech sounds (both auditory icons and earcons) in computer interfaces is an important technology for both sighted and blind users. A prototype hypermedia system for blind users was developed as part of the TIDE ACCESS Project, and initial evaluations showed that users would benefit from and like to see the addition of non-speech sounds, in an attempt to improve orientation, feedback, and correct command usage. The hypermedia prototype which was evaluated at the end of the project included 9 non-speech sounds which complemented the synthetic speech and braille output. Participants found the sounds to be useful while learning the system as well as when they were more experienced, found them easy to remember, and both their observed performance and their subjective responses indicated that they felt well-oriented and were confident in using the correct commands, compared with participants in earlier evaluations using the system without non-speech sounds. Implications for the use of non-speech sounds in other systems (such as WWW browsers and digital talking books) are discussed.

1. Introduction: The Use of Non-Speech Sounds in Interfaces

Auditory enhancements to mainstream computer systems have become more prevalent in recent years in order to compensate for both the information overload of more and more complex visual interfaces, and for the increasing need for eyes-free and often hands-free interactions. Auditory enhancements can include synthetic speech, digitised speech and non-speech sounds. The purposes of auditory enhancements are varied and have many benefits, summarised by Kramer [13]: for example, for eyes-free work, rapid detection, alerting, orienting, backgrounding, parallel listening, acute temporal resolution, affective response, and auditory gestalt formation. When coupled with other interface channels, auditory information has many other benefits including non-intrusive enhancement, increased perceived quality, superior temporal resolution, high dimensionality, and engagement.

[1] Sensory Disabilities Research Unit, University of Hertfordshire, Hatfield, Herts, AL10 9AB, UK.
[2] Computer Sciences Corp., 266 Second Ave, Waltham MA 02154, USA.

Research on the use of non-speech sounds has been dominated by two distinct approaches. Ga developed *auditory icons*, everyday sounds mapped to computer events by analogy with every sound-producing events [8, 9]. These sounds are intended to be easy to learn and remember si they are 'naturalistic', and Gaver included sounds such as scraping sounds for dragging, a pou sound for copying, and a crashing sound for dropping a file in the trashcan. The alternative appro developed by Blattner and colleagues [3] uses a language built from short sequences of tones, i pitches and rhythms, called *earcons*, which are associated with actions and objects. These tona patterns of tones form a grammar for the interface which needs to be learned. It is clear that altho some of Gaver's natural sounds have an obvious appeal (a printer sound for the printing process, example), they are unlikely to be adequate for presenting highly complex interface messages, Edwards [7] and Mynatt [16] discuss. Therefore either a type of Blattner's grammar of sounds i be more appropriate, or a combination of the two approaches.

In this paper we will use the general term *non-speech sound* to refer to any auditory signal whic non-verbal in nature, thus covering both earcons and auditory icons. Non-speech sounds can be to present several different types of information about the interface to the user, for example, st information about an object, confirmation of a command execution, status information about a ci ongoing activity, monitoring of background tasks, and signalling an alarm.

Numerous examples of the successful use of non-speech sounds have been reported in the lite for both sighted and blind users, although their meaningful use in commercial systems is still lir Several examples are presented below to indicate the variety of systems utilising non-speech sou

Earcons have been usefully added to a WWW browser [1] to help sighted users to perform effectively. Real-world auditory cues (e.g. clicks, typewriter, glass breaking) were used to pr feedback for normal data transfer, for the execution of external 'helper' applications, and for conditions, and to indicate other information, for example, the type of link destination (text, gr video, audio or application), the relative file transfer time, and the size of the file. Some sour to be actively learned, but users were able to perform concurrent tasks using the sound cues, able to get a feel for the information at the end of a link, and were happy with the soun confirmation of their actions or events, leading them to predict the results of future actions on the of past interactions. In addition, a framework for the use of non-speech audio in HTML inte has been developed called AHA [11] based on several studies with prototype browsers and users. This framework provides principles for presenting HTML information using non-s sounds to provide users with good information about the structure of a page. This framewo stresses the need for user configuration to suit individual user needs.

Brewster's work on earcons has shown different aspects of improved performance in a number of interfaces. For example, a Macintosh menu with sounds helped mouse users to select command more quickly with reduced subjective workload, and increased recovery from error [5]; a telephone-based menu system with earcons for the different main menu options and different levels of their sub-menus aided orientation, and users did not find the sounds hard to remember [4]. Arons' auditory interface for the Macintosh [2] allows sighted users to skim through digital audio files, and includes non-speech audio to provide terse, yet unobtrusive navigational cues, which are reported to have been effective at providing unobtrusive feedback, avoiding the need for many verbal messages. Although many users were not aware of all the information the sounds were intended to convey, they were using the sounds to aid their task performance. Increased training in the meaning of the sounds is likely to make the sounds even more effective.

Although blind users have long relied on synthetic speech for their interfaces, developers have been slow in realising that there are many instances where the addition of non-speech sounds would make interfaces for these users faster, less verbose, more supportive, and generally more comfortable to use. Speech messages must be processed in a time-consuming serial manner, generally from left-to-right, and we find it difficult to attend to more than one message at a time. Non-speech sounds, on the other hand, can be very short, can contain a great deal of information, can be processed in parallel to speech, and can *replace lengthy verbal strings*.

Several systems have investigated the use of non-speech sounds in interfaces specifically for blind users. For example, Edwards' Soundtrack [7], was a mouse-based system based on earcons to indicate the spatial arrangement of the menu options on the screen. This simple interface was shown to be viable, but was acknowledged to be limited in its usefulness. However, Edwards concludes that sounds should communicate as much information as possible, to reduce the user's memory load when working. Two projects have investigated access to graphical user interfaces such as Microsoft Windows, and both have used non-speech sounds. The GUIB (Graphical User Interfaces for Blind People) Project [10] used a mixture of real-world and musical sounds to indicate objects, properties and events, which improved performance, understanding, confidence and correct usage. The Mercator Project [16, 17] used types of auditory icons with different properties to indicate objects, events, size, state etc. Users in the Mercator evaluations did have to learn the meanings of some of the sounds, since not all events can be easily mapped to a true auditory icon, and thus Mercator also used the metaphorical properties of an event and a sound to convey information.

Other systems which have explored the use of non-speech sounds in interfaces for blind users, either alone or in combination with speech, include the menu system developed by Karshmer and colleagues [12]; the Emacspeak screenreader [20], the AudioGraph system [21]; and the MathTalk system for exploration of mathematical expressions [22].

Much of the existing research investigates the nature of the non-speech sounds; for example
acceptability as examples of sounds for different classes of event or object, or their mean
There are relatively few examples of full evaluations of whole systems or of prototype system
include non-speech sound as an integral part. This paper presents the evaluation of such a sys
Both types of research are necessary, and the results are highly complementary.

2. The ACCESS Hypermedia System

In the context of the TIDE ACCESS Project a hypermedia application was developed. This con
of a web of hypermedia nodes of information presented by synthetic and digitised speech, a Bra
display, audio files and pictures with descriptions, and a specially-designed non-visual inter
making use of digitised speech and non-speech sounds. Details of the full system and eval
results are given elsewhere [15, 19]. Potential blind end-users were involved throughout the
process, and to complement results from full evaluations of the whole system, several small stud
on the use of non-speech sounds in the interface were conducted.

Non-speech sounds were added to the ACCESS hypermedia non-visual interface where
evaluations had shown that performance and usability might be improved with additional feedb
especially non-verbal feedback. The non-speech sounds used in the final prototype were a mixtu
auditory icons and earcon-type sounds, in an attempt to; (a) enhance specific interactions with
system; (b) utilise any appropriate 'realistic' sounds where possible; (c) make use of simple
metaphorical abstract sounds for the remaining situations; and (d) maintain a coherent set of sound

2.1 Demand for Non-Speech Sounds

As part of an extensive user requirements study at the start of the project, a large sample of b
students, their teachers and blind computer experts (181 in total, see Table 1) were surveyed
their preferences for hypermedia systems. All participants were enthusiastic about the use of n
speech sounds to present status information and confirmation, with 74% considering the inclusi
non-speech sounds to be very important (see Table 1), and stating that many systems try to pr
too much information in words which are time-consuming to listen or to read.

Table 1: Number of participants in favour of non-speech sounds

Participants	Number/Total	Percent
Blind Students	65/76	86%
Blind Computer Experts	61/92	66%
Sighted Teachers of Blind	9/13	69%
Total:	135/181	74%

The student participants were significantly more in favour of the use of non-speech sounds (86% favouring their use) than either the computer experts or the teachers (Chi-square=8.27, df=1, p<0.005). One reason for this may be that the older and more experienced participants have used systems which use sounds in ways which are not functional and do not enhance usability (common in many commercial interfaces).

12 Methods of Link Presentation

In the interface of the first prototype synthetic speech was used to present textual information together with a few digitised audio files. Links were indicated by raising the pitch of the reading voice for the link word(s). Seven blind students completed a full evaluation of the prototype using a range of standard tasks. Their performance was observed and recorded and used together with their opinions of the system to assess the usability of the system, and the full evaluation results are reported in [15]. Participants rated various aspects of the interface using a 5-point Likert rating scale (1 = least positive and 5 = most positive).

Participants generally found it easy to identify the link word(s) using the pitch change, giving a mean rating of 4.0 ($\bar{x}$=4.0, n=7)[3]. However, the pitch change sometimes distorted the pronunciation making the word(s) difficult to hear. For this reason, alternative speech parameters (such as a completely different voice), or non-speech sounds were suggested for link presentation. The acceptability of alternative methods for presenting links was investigated, and participants most favoured the use of a sound played throughout the duration of the link word ($\bar{x}$=3.71, n=7), a sound played just before the link word ($\bar{x}$=3.42, n=7), or a sound played before and after the link word ($\bar{x}$=3.14, n=7). Many participants suggested using *both* a sound and a voice change to clearly mark links.

Participants rated the idea of using different sounds to indicate the type of node at the destination of a link, and thought this would be quite useful ($\bar{x}$= 3.57, n=7), so as part of a separate demonstration with another 6 blind students, this idea was tested. Different sounds were played before the link word (spoken in synthetic speech) to indicate: a short definition textual node (*Bing*), a textual node (*Zwoop*), an audio node (*Whistle*), and a picture with description node (*Trill*). The sounds were all short (less than 0.75 sec) and were not meant to have any intuitive mapping with the node they referred to and therefore their meaning had to be learned.

All participants appreciated the type of information the sounds gave, especially to save them from jumping to a node which might not be useful to them. Participants learned and remembered the difference between the sounds very easily, and said they liked the sounds themselves. Users of a

---

[3] $\bar{x}$ represents the mean rating for the given number of participants (n).

Braille display heard the sounds when they loaded the node, so they knew even before reading the display that there were several different links in the node, and they found this useful. Several participants suggested that the sounds could indicate additional attributes of a link, for example, the size of the destination node, or whether the node had been visited before, although some participants suggested merely changing additional parameters of the synthetic speech for the link word.

2.3 Event Confirmation and Information

During the evaluation of the first prototype, only one sound *(Uh-Oh)* was implemented in the interface, to indicate a 'dead-end', at the end of a node, or when users tried to use commands which were not available. This was rated as being very useful ($\bar{x}$=4.86, n=7) in providing short immediate information without recourse to a verbal message, and it also made the students laugh - and this helped to maintain their motivation to use the system. However, the sound also seemed to have negative connotations for some participants, making them feel as if they had made a mistake and these participants suggested a less negative sound - such as bouncing off an elastic wall (e.g. *Boing*).

The evaluation also highlighted several places where the addition of non-speech sounds could improve users' interactions. Participants were observed to be confusing certain commands and situations and thus becoming disoriented. This was due to a lack of confirmation of commands and information about the status of what users were reading. The two main problems were: (a) failure to distinguish between commands for (i) reading within a node (i.e. read next/previous paragraph using the Read Right/Left commands) and (ii) jumping to a new node (i.e. selecting a link, or using the Back/Forwards commands); and (b) failing to distinguish between (i) arriving at the heading of a new node (by selecting a link, or using Back/Forwards) and (ii) arriving at the heading of the current node by simply jumping up to the current node heading (using the Heading command).

The interface of the second prototype included 9 non-speech sounds to address these problems, providing information about the state of information, and to confirm the results of a command execution. Non-speech sounds were only added where there was a perceived necessity, rather than adding sounds to the whole interface. In addition, the sounds were designed to appeal to and engage the blind student participants and therefore included a mixture of auditory icons where appropriate and simple earcon-type sounds. For example, a *Paper Riffle* sound (like pages turning in a book) was used to indicate reading within a node, whereas the electronic sliding sound *Zoom* indicated jumping to a completely new node using Back/Forwards. A *Fanfare* indicated arriving at the heading of a new node, and a sliding sound *Swoosh* to indicate sliding up the current page to the heading. A futuristic sound *Transporter* was used to represent jumping to the home node, reminiscent of the Star Trek™ device bringing crew back to the ship from anywhere (see [6] for more extensive use of Star Trek™ sounds in an interface).

A full list of commands and events is given elsewhere [15], and Table 2 shows just the commands and events with which sounds were associated. Links were presented using both a non-speech sound before the link word(s) *and* the link word(s) itself spoken in a different voice, to evaluate this double method of link presentation - which many participants had themselves suggested.

Nineteen blind participants (9 students and 10 adults) completed an extensive evaluation of the second prototype, involving individual use over an extended period, followed by several standard tasks requiring full realistic use of the system, for which objective performance measures were taken. These were complemented by a usability questionnaire-based interview. The full evaluation results are given elsewhere [15] and the details of the results for the non-speech sounds are reported here.

2.3.1 Ease of Use and Usefulness

Participants' performance and responses showed clearly that the system was easier to use than the first prototype, and participants were observed to be reacting to the non-speech sounds quite confidently although they were not consciously directing their attention to the sounds. They also appeared to be more relaxed and secure when working with the system, worked faster and made fewer errors, and reported that they felt well-oriented and knew which commands to use. Participants commented that they were receiving a lot of information quickly and unobtrusively, without having to execute additional commands or wait for long verbal messages. Most of these participants had no prior experience with the system, and therefore this level of confidence amongst novices is encouraging.

The Usefulness of non-speech sounds was rated as very high initially (i.e. when learning to use the system) ($\bar{x}$=4.38, n=19), becoming only slightly less useful when participants became more experienced with the system ($\bar{x}$=4.25, n=19). For most participants, the sounds were more useful as information and confirmation while learning the system, whereas for a few, the sounds were initially confusing, and became more useful with experience. Even though many of the sounds used might seem lighthearted, none were reported to be irritating or distracting, although some were reported to be too long and too loud. The sounds were especially valuable to users of Braille-only output, who were able to assess their location using the sounds without having to spend time reading the display.

2.3.2 Memorability and Identification

To assess the memorability of the sounds, participants were played each sound in turn and were asked to name the event or command to which it belonged, and to rate how easy each sound was to remember. The numbers of correctly identified commands and ratings for memorability are shown in Table 2. Most participants were able to correctly name two-thirds of the sounds, although one sound was not heard often enough to be learned (the sound for the Heading command).

Table 2: Numbers of correctly identified sounds and ratings for ease of remembering

Command/Event	Non-Speech Sound	Participants Correctly Identifying Sound	Mean Rating for Ease of Remembering 1=low - 5=high
Read Left/Right (next/prev paragraph)	Paper Riffle	19	4.92
Dead-End	Donk	19	4.71
Back	Upwards Slide Zoom	17	4.38
Link	Ding	16	4.31
Heading	Fanfare	17	4.17
Centre-Point	Pop	18	4.00
Home	Transporter	18	3.94
Forwards	Downwards Slide Zoom	11	3.52
Heading Command	Swoosh	0	-
Overall Mean Memorability Rating for 8 Remembered Sounds:			4.24

Two of the nineteen participants were unable to identify many of the sounds, although most participants said they had no trouble identifying the meaning of the sounds while working in the system. Most of the commands were rated as being easy to remember (overall $\bar{x}$= 4.24, n=19 for the 8 identified sounds).

3. Conclusions

The non-visual hypermedia interface developed as part of the ACCESS Project has shown that the use of non-speech sounds for informing users about commands and events made it easier for users to work with the prototype system, providing them with short, unobtrusive feedback about the results of their actions and the status of the information in the hypermedia material. Some of the important implications are presented below:

- Non-speech sounds have been shown to be a valuable addition for both speech-output and Braille output systems, to provide immediate and short state information and command confirmation.
- There seems to be a good rationale for using a mixture of both auditory icons and earcons especially when users might require special motivation to use a system (e.g. students). For instances when an auditory icon is especially appropriate (such as a dead-end represented by a bang or a bounce against a wall; or going to the Home node represented by a Star Trek™ type transporter) it would be a pity not to use one. For other events, such as indicating the nature of a link (e.g. its destination node type, size, whether it has been previously followed etc) or providing an auditory overview of the contents of a page, earcon-like sounds seem more appropriate since they can form part of a coherent grammar of sounds.
- The choice of sounds should be appropriate to the event, and piloted with a range of users for their success at conveying the desired information to ensure that the selected sounds are effective.

- Care should be also taken to ensure that the selected set of non-speech sounds are consistent - in terms of length, volume, and placement with a command or event, so that the user is always clear as to the association of a sound to a particular command/event.

- Users have their own distinct preferences for non-speech sounds while they are learning and then using a system, so a highly flexible user-configuration component should be included for users to specify; (a) the exact sound for each event/command, (b) whether each specific sound is on/off, or (c) whether all sounds are on/off. This configuration should be easily accessible.

- The use of any non-speech sound is likely to require some training and practice. As Brewster [4] discusses for his telephone-based menuing system, a mixture of documentation and on-line training with examples could be used to give users the correct interpretation for the sounds in a system. This will help to ensure that users have the intended model for the sounds.

- The number of sounds used should be carefully considered. Participants in these studies liked the use of these few sounds for the most important or difficult events/commands, but did not wish the system to be full of non-speech sounds, which might become irritating. In addition, there is likely to be a threshold for learning and remembering of many sounds. For example Patterson [18] found that naive listeners of aircraft warning sounds could learn up to 7 warning sounds reasonably quickly, but beyond that, learning was slow but steady (and only up to 10 distinct signals). In addition, individual differences in the ability to learn and remember only 9 sounds were revealed here, so care should be taken to make very efficient use of added non-speech sounds.

- Further experiments will need to address the difference of context-dependent sound use and isolated sound testing, since our participants were observed to be using the information contained in the non-speech sound correctly while using the system, but some had trouble naming the same event/command when the sounds were played in isolation.

- These experiences should help designers of other non-visual systems to include non-speech sound in valuable ways, and recommendations have already been proposed for other similar systems, such as Digital Talking Books [14]. One mainstream system already using sounds to improve feedback to users is Windows 95, which now includes a set of 'soundschemes' providing a range of different sounds for different events, which can be tailored to suit individual users' needs and preferences.

4 Acknowledgements

We would like to thank all our enthusiastic participants, and our partners in the ACCESS Project (TIDE Project 1001). Sarah Morley is sponsored by a postgraduate studentship award (number S00429537025) from the Economic and Social Research Council of the U.K.

5. References

[1] ALBERS, M. Auditory Cues for Browsing, Surfing, and Navigating the WWW: The Audible Web. Procee ICAD'96: Third Annual Conference on Auditory Display. Reading, MA: Addison Wesley, 85-90. 1996.

[2] ARONS, B. SpeechSkimmer: A System for Interactively Skimming Recorded Speech. ACM Transactions on Computer-Human Interaction, 4 (1), 3-38. 1997.

[3] BLATTNER, M.M., SUMIKAWA, D.A. and GREENBERG, R.M. Earcons and Icons: Their Structure and Common Design Principles. Human Computer Interaction, 4, 11-44. 1989

[4] BREWSTER, S. Navigating Telephone-Based Interfaces with Earcons. In H.Thimbleby, B.O'Conaill, P.Tho (eds) People and Computers XII, Proceedings of HCI'97. London: Springer, 39-56. 1997.

[5] BREWSTER, S.A. and CREASE, M.G. Making Menus Musical. In S.Howard, J.Hammond, and G.Lingaard Proceedings of Human-Computer Interaction: INTERACT'97,,389-396. 1997.

[6] COHEN, J. "Kirk Here:" Using Genre Sounds to Monitor Background Activity. In S.Ashlund, K.Mullet, A.Henderson, E.Hollnagel and T.White (eds) Adjunct Proceedings of INTERCHI'93, 63-64. 1993.

[7] EDWARDS. A.D.N. Soundtrack: An Auditory Interface for Blind Users. Human Computer Interaction, 4,45 1989.

[8] GAVER, W.W. The SonicFinder: An Interface That Uses Auditory Icons. Human Computer Interaction, 4,67 1989.

[9] GAVER, W.W. Using and Creating Auditory Icons. In G. Kramer (ed) Auditory Display: Sonification, Audification and Auditory Interfaces. Reading MA: Addison-Wesley, 417-446. 1994.

[10] GUIB Project Final Report: Textual and Graphical User Interfaces for Blind People. London: RNIB. 1995.

[11] JAMES, F. Lessons from Developing Audio HTML Interfaces. Proceedings of ASSETS '98: Third Annual AC Conference on Assistive Technologies. New York: ACM Press, 27-34. 1998.

[12] KARSHMER, A.L, BRAWNER, P. AND REISWIG, G. An Experimental Sound-Based Hierarchical Menu Navigation System for Visually Handicapped Use of Graphical User Interfaces. In Proceedings of ASSETS '94 First Annual ACM Conference on Assistive Technologies. New York: ACM Press, 123-128. 1994.

[13] KRAMER, G. An Introduction to Auditory Display. In G. Kramer (ed) Auditory Display: Sonification, Audification and Auditory Interfaces. Reading MA: Addison-Wesley, 1-77. 1994.

[14] MORLEY, S. Digital Talking Books on a PC: A Usability Evaluation of the Prototype DAISY Playback Software. In Proceedings of ASSETS '98: Third Annual ACM Conference on Assistive Technologies. New Y ACM Press, 157-164. 1998.

[15] MORLEY, S., PETRIE, H., O'NEILL, A-M. and MCNALLY, P. Auditory Navigation in Hyperspace: Desi and Evaluation of a Non-Visual Hypermedia System for Blind Users. Proceedings of ASSETS '98: Third Ann ACM Conference on Assistive Technologies. New York: ACM Press, 100-107. 1998.

[16] MYNATT, E.D. Transforming Graphical Interfaces Into Auditory Interfaces for Blind Users. Human Computer Interaction, 12, 7-45. 1997.

[17] MYNATT, E.D. Designing Auditory Cues for Mercator. In S.Howard, J.Hammond, and G.Lingaard (eds) Proceedings of Human-Computer Interaction: INTERACT'97, 236-243. 1997.

[18] PATTERSON, R.D. Guidelines for Auditory Warning Systems on Civil Aircraft. Civil Aviation Authority London. 1982.

[19] PETRIE, H., MORLEY, S., MCNALLY, P., O'NEILL, A-M., and MAJOE, D. Initial Design and Evaluat an Interface to Hypermedia Systems for Blind Users. Proceedings of Hypertext 97. New York: ACM Press, 4 1997.

[20] RAMAN, T.V. Auditory User Interfaces. Toward the Speaking Computer. Boston: Kluwer Academic. 1997.

[21] RIGAS, D.I. and ALTY, J.L. The Use of Music in a Graphical Interface for the Visually Impaired. In S.Ho J.Hammond, and G.Lingaard (eds) Proceedings of Human-Computer Interaction: INTERACT'97, 228-235. 199

[22] STEVENS, R.D., EDWARDS, A.D.N. and HARLING, P.A. Access to Mathematics for Visually Disabled Students Through Multimodal Interaction. Human Computer Interaction, 12, 47-92. 1997.

THE ARCHITECTURE OF THE BLIND USER ORIENTED HYPERTEXT AUDIS

Ivan Kopeček

Faculty of Informatics
Masaryk University
Botanická 68a, 602 00 Brno, Czech Republic
e-mail: kopeeek@fi.muni.cz
http://www.fi.muni.cz/~kopecek/

Abstract

The architecture of the blind user oriented hypertext AUDIS developed at the Faculty of Informatics, Masaryk University Brno, is described in the paper. AUDIS is developed primarily as a support, which would help visually impaired students to study various materials. It can also serve as an Internet browser.

1. Introduction

Speech synthesizer and "screen access" ("screen reader") software represent basic utilities used by blind users to obtain hypertext information. Although present conventional hypertext and fulltext systems are very sophisticated and user-friendly, they are not very convenient when they are to be used by visually impaired people. The reason is the graphical interface and absence of the features fulfilling special needs of the blind people ([1]).

In what follows we will describe the basic features of the hypertext system AUDIS developed for blind users. The hypertext system AUDIS is primarily developed to provide comfortable accessibility to textbooks for visually impaired students. It maximally pays attention to the special needs of users and allows for comfortable customization. The system is supported by the utilities for text conversion and rendering. Next versions of AUDIS will allow the system to be used in World Wide Web browser as well.

Even though AUDIS is primarily developed for visually impaired people, the architecture of the system is designed in such a way that the system will be convenient even to a larger group of people with special needs, e.g. for print impaired people.

2. Control of the System

The user controls the system by means of speech commands combined with keyboard commands. This access optimizes the effectiveness of the control by joining advantages of both methods. Every command exists in both versions, i.e. in speech command version and in keyboard key version. This is due to possible application of the system for print impaired people.

Speech commands consist of the key words and are recognized by built-in speech recognizer. The system enables customization of the commands by renaming key words or by defining alternatives.

To use the basic speech commands the recognizer need not be trained, it is automatically customized to the user voice. However, some simple training of the recognizer increases reliability of the speech recognizer when larger number of speech commands is used.

Sound is the main output of the system. It can be used in the form of:

- Synthesized voice produced by the syllable based speech synthesizer DEMOSTHENES (see e.g. [6]). This type of sound output is used for reading hypertext data. The used speech synthesizer applies basic prosodic features based on corpora analysis (see e.g. [7]) and enables use of various voice types that can be configured by the user. Various voice types are used to distinguish various types of information (e.g. normal text, underlined text, links etc.).
- Sampled voice, which is used for all standard messages to the user. Even sampled voice output can be configured. Various types of sampled voices are used to help the user to distinguish various types of messages.
- Sound generated by the sound synthesizer, MIDI, wave tables or special samples. This type of non-speech sound will be used first for environmental sounds applied to provide feedback to user actions ([2]), secondly it can be used for earcons (non-speech glances used to give the user an overview by listening, ([8]).

3. Hypertext Data Structure

The following important features influence the data structure;

- Maximal HTML compatibility. This is not only due to Internet but also because of general compatibility of data.
- Maximal possibilities for customization of the system and data.
- The primary goal of the system is to provide comfortable accessibility to students textbook material for visually impaired students - the possible use of AUDIS as a Internet browser has secondary meaning.

There are the following types of data files that are used by AUDIS:

- HTML data; more precisely, the system will accept only a defined subset of HTML (depending on the version of the system). Constructions and expressions outside the defined subset will be ignored.
- System configuration files are used for general customization of the system.
- Data customization files are used for user customization with regard to the data (text information included in the hypertext).
- Task definition data. It is useful to provide the system with the information whether it is used as a local hypertext system supporting textbook material or as an Internet browser.

4. Data Preparation

Hypertext data should be carefully prepared to be used by the visually impaired students in accordance with the system data structure and with the basic features of the system (see Section 6). Although a simple structure based on the defined subset of HTML is acceptable for the system, some other text should be added to the basic text and some additional pieces of information should be included to make the data more comfortable for the visually impaired users.

This includes;

- Global modification of the text. The number of the pieces of information that are not conv for sight impaired (graphics, mathematical expressions, tables etc.) should be reduced and/or explained, if possible, by means of textual form ([3]). Shape of the text should be modifi possible, respecting the users needs.
- Rendering and interpreting of graphics, tables, mathematical expressions and other non-textual information.
- Defining underlined sentences (sentences containing important information will be phoneti accentuated) and other meta-information in accordance with the data structure convention
- Adding summaries in accordance to AUDIS hypertext data structure conve (see Section 7).
- Reducing the number of links and references, if possible.
- Building the tree structure of the system from basic linear structure; this is supported by automatic generation of the optimal tree structure (see the next Section 5).

Some other pieces of information will be generated automatically (e.g. earcons, sound g environmental sounds).

5. Decision Tree Construction

Although the information structured by hypertext links is convenient for visually impaired pe too frequent links may decrease lucidity of the text.

To make the structure of the system as simple as possible we use structure of hypertext sy consisting of menus (menu pages) and chapters (text pages). Menus contain a title and links and form the decision structure of the system. Chapters are linearly ordered elements of the The links included in the chapters should be used only inside the sequence of the chapte should not be used to address menus).

Such a structure is more transparent and lucid for visually impaired users than a gene structure. Formally such structures can be represented by rooted trees (root of the tree is the menu of the system, inner nodes are menus and leaves correspond to chapters).

When a hypertext structure is created from the basic linear text items we have to define the structure of menus of the decision tree. It is clear that it should be done in such a way that the resulted structure is in some sense optimal from the users point of view. The optimality criterion is motivated by the requirement of the minimal average number of menu items that must be read by the system and evaluated by the user to find the required text page.

The problem of finding optimal rooted trees for such optimality criterion is solved in ([5]). Let us remark that in the quoted paper the criterion is formulated for another problem with different motivation, but it leads to the same formalization as in the present case. We will briefly recall the main relevant result.

We consider all sets to be finite (card(m) denotes the number of elements of a set m). Rooted trees are denoted by ordered pairs R = (V, A) of vertices (nodes) and arcs (edges). We use also notation V(R) for R and A(R) for A. A root is denoted by r. We denote the set of all successors of a node x by q(x) and by L(R) the set of all leaves of the tree R.

Let R = (V, A) be a rooted tree, x, y $\in$ V, and suppose that there exists an (n+1)-tuple (x_0, x_1, ..., x_n) such that (x_{i-1}, x_i) $\in$ A, x_0 = x, x_n = y. The set {x_0, x_1, ..., x_n} will be denoted P(x,y) and we use notation y < x. If x_0 = x = r then n is called level of y and is denoted lvl(y). Further, lvl(R) will denote the number max {lvl(x); x $\in$ V(R)}.

Definition of Optimality. Let R = (V, A) be a rooted tree, x, y $\in$ V, card(V) > 1 and y < x.

We put $w(x,y) = \sum_{z \in P(x,y)-\{y\}} card(q(z))$ and $W(R) = \sum_{z \in L(R)} w(r,z)$;

The function W represents optimality measure. We shall say that a rooted tree R = (V, A) is W-optimal, if for each rooted tree R' satisfying L(R) = L(R') the condition W(R) $\leq$ W(R') holds;

Definition. A rooted tree R = (V, A) is called L_3 tree if it satisfies:

(1) card(q(x)) < 4 holds for every x $\in$ V,

(2) card(q(x)) < 3 implies q(x) $\subseteq$ L(R).

Balanced quasi-ternary tree is an L_3 tree that satisfies:

(1) lvl(x) < lvl(R) - 1 implies card(q(x)) = 3;

(2) lvl(x) = lvl(y) = lvl(R) - 1 implies |card(q(x)) - card(q(y))| < 3;

(3) card(q(x)) $\neq$ 1 for every x $\in$ V.

Theorem ([5]). Balanced quasi-ternary trees are W-optimal.

This result makes it possible to propose the decision structure of the hypertext in an automatic and optimal way (in the form of balanced quasi-ternary trees).

6. Basic Features of the System

In this section we will present the basic principles on which the system is built from the user poi of view.

- The system enables comfortable control by means of combination of speech commands and ho-key commands. Speech commands are supported by speech command thesaurus that allows u express a command in several ways, making the control of the system more intuitive.
- Easy customization and configuration are important features of the system. This is related to the control commands, mode and type of speech synthesis output, graphic output for print impai people with sight, information data structure, and other properties and options of the system
- It is important to enable the user to obtain the information quickly and to get always informational overview. This is supported first by various output speech modes and outp speech rates, second by speech summaries, audio glances, earcons and environmental sound
- The orientation of the user is supported by the information about the position that is alw accessible in speech form as well as in the form of audio glances, earcons and environm sounds.
- User's speech remarks can be assigned to each information unit (menu, chapter, sect sentence). Each information unit can be underlined or masked. The user and system informati attributes (masking, underlining, speech mode and voice types etc.) are always saved and can always reset.
- The system is provided with tree, linear and fulltext type of searching and with the datab speech bookmarks. Summaries of the related information unit are always associated to bookmark and can be used for orientation.

- Control of the speech output is fully supported. Each information unit can be skipped or repeated, the speech output can be stopped, speeded up or slowed down.

p. Information Units of the System

There are four basic information units of the system (here we suppose the data in the form of textbook): menus, chapters, sections and sentences. Menus form tree decision structure (its creation during the stage of data preparation is supported by utilities that yield optimal decision tree for given number of chapters), linearly ordered chapters are leaves.

Chapters are divided into sections. Menus, chapters and sections must be provided with summaries that consist of the text part and audio glances (earcons). The non-text parts of the summaries are generated automatically. The following scheme corresponds to the chapter structure.

Title of the chapter

Summary of the chapter

Title of the section

Summary of the section

Text of the section

Title of the section

Summary of the section

Text of text section

The user may choose whether some types of summaries and audio glances will be masked (then they are not outputted automatically).

Sentences can be non-underlined, underlined, masked and non-masked (in all combinations). T
user can overwrite the original attributes (as they were assigned in the time of data preparation)!
the original values can be always reset. User-defined attributes can be also saved and reset.

8. Conclusions

We have presented the basic architecture principles and features of the developed blind
oriented hypertext system AUDIS. The system is developed and tested in close co-operation wi
the Czech Union of Blind and Visually Impaired. The further work will be oriented to
implementation of basic modules, creation of the data preparation programming support
developing methods for rendering non-text data into speech/sound output.

Acknowledgements

The author is grateful to Hana Bubeníčková (Union of Blind and Visually Impaired) and Karel P
(Faculty of Informatics, Masaryk University) for reading a draft of this paper, for valu
comments, and helpful ideas. The development of AUDIS is supported by the Grant of the Cz
Ministry of Education No. VS 97028.

References

[1] BUBENÍČKOVÁ, H., KOPEČEK, I. (1997) Informacni systemy pro nevidome (Information Systems for Blind People, in Czech). Proceedings of "URBIS '97", Brno, April 1997, 34-36.

[2] DARVISHI, A. (1996) World Wide Web Access for Blind People: Problems, Available Solutions and an App
for Using Environmental Sounds. Proceedings of ICCHP'96, Linz, 1996, pp. 369-373.

[3] KAHLISH, T (1996) Improving Access to Hypertext Based Study Material for the Blind. Proceedings of ICCHP'96, Linz, 1996, pp. 229-236

[4] KOPEČEK, I. (1997) Speech Synthesis of Czech Language in Time Domain and Applications for Visually Impaired. Proceedings of the 2nd SQEL Workshop on Multi-Lingual Information Retrieval Dialogues, Pilsen, April 1997, pp. 141-145.

[5] KOPEČEK, I. (1997) Optimal Trees for Searching in Codebook. Proceedings of SOFSEM'97, Lecture Notes in Computer Science, Springer, 1997, pp. 456-464.

[6] KOPEČEK, I. (1997) Syllable Based Speech Synthesis. Proceedings of the 2nd International Workshop Speech and Computer- SPECOM'97, Cluj-Napoca, 1997, pp. 161-165.

[7] PALA, K., RYCHLÝ, P., SMRŽ, P. (1997) DESAM – Annotated Corpus for Czech. Proceedings of SOFSEM'97, Lecture Notes in Computer Science, Springer, 1997, pp. 523-530.

[8] STEVENS, R.D., WRIGHT, P.C., EDWARDS, A.D.N., BREWSTER, S.A. (1996) An Audio Glance at Syntactic Structure Based on Spoken Form. Proceedings of ICCHP'96, Linz, 1996, pp. 627-635.

How Computers in the Czech Republic Help Blind Persons to Overcome Information Barrier and its Adversary Impact

Hana Bubenickova

Historic summary:

- *till 1989 : Picht's brailler (typewriter for blind),*
  CCTVs (close circuit TV) for severally visually impaired persons,
- *December of 3rd 1989 : foundation of Czech Blind United (CBU),*
- *1990 : Import Eureka A4 - a non-standard PC option intended for the blind (400 users in Czech republic, 200 users in Slovak republic),*
- *1991 : The programs LP-DOS, Magic in MS-DOS for severally visually impaired persons,*
- *1992 : CBU founded a Digitizing Centre in Prague,*
- *1993 : Other digitizing centres were founded in Czech republic,*
  Start of operation and services of BBS BraillNet (+420 2 66038528 - 9) ,
  Czech voice response in MS-DOS to standard PC - program „KUK" for blind,
- *1994 : Using of programs LP-DOS, Magic in Windows 3.11,*
  Using of Recognita Auge, Braille displays ALVA, CombiBraille
  Start of testing, evaluating, advisory and based training in area of access technology
- *1995 : Braille display Thymus (Czech product), PC fitted with relevant options from blind and severally visually impaired persons - 100% contribution of a social benefit*
- *1996 : Czech voice response in Windows (Windows 3.11 and Windows 95) called „WinTalker", Digitizing centres have transformed in National Adaptive Technology Centres and National Information & BraillNet Centre CBU,*
  Internet server BraillNet Plus starts its operation (http://www.braillnet.cz)
- *1997 : Internet for blind and any disabled person in CR.*

The blind in our country has already the possibility to use computers fitted with sp... options for several years. *Czech Blind United* (CBU), the independent non-profit ... organisation of the blind, founded in 1989, is responsible for proliferation of these extras withi... thus, its step by step activities help to reduce the above-mentioned information barrier. In 199... CBU founded a Digitizing Centre in Prague aimed at electronic processing of printing materials i... adaptation of digitized texts that can be read using so called *Eureka A4* - a non-standard PC op... intended for the blind. *BBS BraillNet* is intended for storing and retrieving various texts; currently it offers not only the library containing books in a digitized form but also various text... magazines, newspapers, journals etc.; it offers also electronic mail and other standard BBS servic... for example application gateway, that enables opening of various applications on a distant comp...

The non-standard form of Eureka and its incompatibility with other standard IBM PC r... in development of so called *voice response* that can be used in standard PC under MS-DOS. T... most used voice response in MS DOS is a program called „KUK". Implementation of voice resp... in Windows was supported by the program called CS Voice; originally, it was intended only...

reading text files in Windows. Up to now, its author has created a very stable and reliable voice response called Wintalker; it can be used under Windows 95 and is able to communicate within Internet text environment.

Persons suffering from severe visual debility can use high quality imported *programs enabling image enlargement* (LP-DOS, Magic); they can use MS-DOS, Windows 3.11 as well as Windows 95. Specially created programs „SCAS" and „Magna Vista" working on a scanner can enlarge a printed document which enables to handicapped persons to read any printed text.

Recognita Auge and Recognita R3 *programs make printed text readable for the blind* in our country; they are imported from Hungary and both their menu and code pages are adapted to our national environment.

Our organization also co-operates with *Faculty of Informatics at Masaryk University, Brno*. Here, an advanced type of voice synthesizer is being developed (see: Kopecek I. Syllable Based Speech Synthesis, Proceedings of SPECOM'97, Cluj Napoca 1997 pp. 161-165). Other interesting activities, e.g. development of a navigational device for the blind using a computer and communications engineering or development of a blind oriented hypertext system, are being also topics of research at this faculty.

Computers that are adequately configured and fitted with all necessary options and extras enabling persons suffering from severe visual debility their effectively use are still *rather expensive in our country. Braille displays* - touch outputs of personal computers - are also too expensive for an ordinary Czech blind person. Unfortunately, students or mentally working people who have to process a lot of texts, e. g. lawyers or interpreters, are currently the only ones who can get the Braille display in our country. Thanks to favourable legislation promoted also by our organisation, the blind and severally visually impaired persons have nowadays the opportunity to buy PC fitted with relevant options because they can get up to *100% contribution* in a form of *a social benefit*. It is estimated that there are more than 1 000 blind computer users and approximately 100 Braille displays users; it should be mentioned that Braille display price is several time higher than that one of a common PC fitted with a scanner and necessary special software and hardware.

BraillNet BBS began its operation in 1993 and currently it is the provider of *BraillNet Plus* Internet server. It offers the same functions as BBS services, simultaneously extends possibilities of visually impaired persons and reduces operation costs all-over the territory. We succeed in connecting the visually impaired persons to Internet under absolutely exceptional and financially very advantageous conditions in a place where they live or work. We are able to offer full scope of services to any person suffering from other disabilities under the same conditions; it means that we have created a real basis on which *the information system intended for any disabled person* in our country can be built. In future, our server should contain the following information:

- List of non-profit humanitarian organisations offering care of handicapped persons and providing services for them;
- List of rehabilitation, compensation, prosthetic, orthopaedic and optic aids;
- Programs aimed at education, self-realisation, job, rehabilitation and re-qualification of the handicapped persons.

Our server is a mean that can help to remove communication and information barrier in our country and assist to take full advantage of new techniques that are able to improve life of many disabled citizens in our country.

RSDr. Hana Bubenickova (E-mail: bubenickova@braillnet.cz),
Director of National Information & BraillNet Centre CBU CR, Na Harfe 9, 190 05 PRAHA, Czech republic
Prepared by: Mgr. Sarka Cernohlavkova (E-mail: cernohlavkova@braillnet.cz)
Regional Information & BraillNet Centre CBU CR, Chaloupkova 7, 612 00 BRNO, Czech republic

STATE OF THE ART CONCERNING THE USE OF COMPUTERS BY BLIND MUSICIANS IN EUROPE AND THE PLAY PROJECT SOLUTIONS.

Didier Langolff, Nadine Jessel, Monique Truquet (UPS, France)

This inquiry was realised with the help of all the PLAY project partners.

Institut de Recherche en Informatique de Toulouse
Université Paul Sabatier
E-mail: langolff@irit.fr

POSTER OVERVIEW

PLAY Project aims to answer the needs of blind musicians. The first part of the project is a careful analysis of user's needs and their interests. This task was carried out as a survey. The poster presents different schemes concerning the results of this inquiry carried out by the PLAY Project Consortium in the seven European countries and the solutions proposed by the PLAY Project.

In a first part, is presented the number of questionnaires sent to seven European countries (which diffused them) and the number of answers.

In a second part, three schemes present the respondents features of the European countries involved in the PLAY consortium:

- Visual impairment.
- Age distribution.
- Musical experience.

The third part describes the computer use by blind persons. Two different schemes are presented and allow to make a comparison between the computer use in everyday life and in the musical domain. The reasons of the few use of computers in the musical domain are described through these schemes.

At the end of the poster, we present the wishes of the informants concerning future system developments. To conclude, a scheme presents an overview of the PLAY project aims to develop a workstation providing the different possibilities which blind musicians are seeking.

PRESENTED WORK

With the increased development of technologies in computer science and communication we have many possibilities of improving the quality of life for disabled persons through the adaptation of these new technologies to their specific needs. The PLAY project tries to answer the needs of blind musicians for access to musical documents.

We will present the state of the art concerning the use of computers by blind musicians with the answers analysis of questions among seven European countries (Denmark, Finland, France, Germany, Italy, Netherlands and Spain)
The big majority of those who answered were blind, some of them were partially sighted and a little part were sighted expert.
43% of them were 40-60 years old. Among them, a big part were professionals in the musical domain what is interesting.
Concerning the use of a computer in the musical domain, we note a maximum of 25% of blind musicians in the Netherlands and only 5% in Spain, against an average of 78% in everyday life. Why so few users? For many reasons:

- **Lack of skills and lack of training.**
- Difficulty of using music programs (use of mouse , too many graphics...).
- Lack of information concerning suitable programs and tools.
- Inadequate technical documents.
- Too old to learn new skills.

Some of them answered that they want to use also computers to access Internet and databases.

Concerning the PLAY project they expressed the following wishes:

- To develop a user **friendly** and **cheaper** system for editing music.
- To scan and translate scores into Braille.
- To print out both ink-print and Braille scores.
- To have the possibility of converting music Braille to Midi and/or printed music, and vice versa which would greatly improve the availability of musical material in Braille.
- To have the possibility of scanning Braille. It would be useful mainly for blind composers who have no possibility of obtaining their scores in a printed form.
- To follow a training in order to be able to use the complete system.
- To have a workstation compatible with the main products used by the blind users.
- To be able to listen to music that is being read.
- To be able to have contacts with other musicians, blind or sighted (email concept).
- To be able to read sheet music, consult catalogues and download items (internet access).

The objective of the PLAY project is to develop a „ PLAY station " managed by an editor which will control the different modules linked together and will allow the user to :

- ➢ Input/output musical text in Braille, or communicating with a commercial program through ETF, NIFF or Midi Files.
- ➢ Save scores in the internal code format (Play Code).
- ➢ Print into Braille.
- ➢ Access catalogues and musical archives through an Internet interface.

The PLAY project will provide a booklet concerning the Internet access. Moreover a study on feasibility of two musical score recognition systems (optical devices) is undertaken.

Education

4
W

t

MathPlus ToolBox, a computer application for learning basic math skills

C. GARDNER

Linn-Benton Community College

630 N.W. 7th Street

Corvallis, OR 97330

ckg@dots.physics.orst.edu

R. LUNDQUIST

Oregon State University

Physics Department

Science Access Project

randyl@dots.physics.orst.edu

Abstract

The MathPlus ToolBox, a computer application, has been designed through a collaborative effort of Carolyn Gardner and Randy Lundquist. The goal of the project was to improve accessibility to math for students with disabilities.

Though there are many computer products which are claim to help students learn math, many do not allow universal access. Often these products don't teach students the mathematical processes. The MathPlus ToolBox uses graphics, sound and captioning so that all students have access to the concepts being taught. There is a variety of tools to help a student learn and practice addition, subtraction, multiplication, division, and fractions

Title of book. Name of Eds.

1 INTRODUCTION

The MathPlusToolBox has been developed in cooperation by Dr. Randy Lundquist with the Oregon State University Science Access Project and Carolyn Gardner at Linn-Benton Community College. It is being developed because of the scarcity of mathematical materials available for students with disabilities. Computers often enable these students to have access to information that is otherwise not available. Though this computer application was developed to help students with disabilities, its universal design makes it appeal to all students learning math.

In this paper, some background will be given about the reasons why the MathPlusToolBox was developed. The tools included in the ToolBox will be described along with some of the reasoning that was behind the design. The development process will be explained along with a look at our plans for the future.

2 BACKGROUND

Persons with disabilities have traditionally been underrepresented in the fields of science, mathematics, engineering, and technology. The future needs of the nation's scientific workforce demand that no group be excluded from participation in education and training that lead to productive careers in the sciences. The National Science Foundation promotes the full inclusion and participation of persons with disabilities in academic studies leading to careers in science and engineering through its Program for Persons with Disabilities. (SRI, 1997) This division of National Science Foundation has funded the development of programs and technologies, and fostered ideas which have improved education for all students.

At this current time, it is especially important not to exclude any group of people from the technical work force. There are jobs which pay well in technology related fields. American companies are unable to find enough qualified people to fill their positions. People with disabilities are often shut out of the job market because of their inability to study mathematics.

One of the reasons that students with disabilities have not studied mathematics is the difficulty of preparing materials for them. This has been particularly true for students who are visually impaired and students with learning disabilities. For the past several years Mrs. Gardner has been researching the uses of technology to teach mathematics to students with learning disabilities. (Gardner, C., 1996) The Oregon State University Science Access Project is working on providing access to information for students with visual impairments. (Gardner, J., 1997) When the project presented several of its research projects at the international conference on Technology and Persons with Disabilities, Los Angeles in March 1997, one of these programs, Triangle, attracted many students, parents and doctors who were looking for a computer program to do basic mathematics. The MathPlusToolBox will help fill that need.

B BACKGROUND BEHIND THE DEVELOPMENT OF THE MATHPLUS TOOLBOX

At the beginning of this project, it was planned to use the sound capabilities of AsTeR (Raman, 1997) to provide access to science textbooks for students who had learning disabilities. This approach was abandoned because of the technological difficulties of combining AsTeR"s speaking capabilities with the visual image of mathematical text on the screen, and because the current mathematical textbooks are too highly dependent on graphical images to explain the concepts. The best approach seemed to be to develop a program which could use the computer capabilities to teach the same concepts being presented in textbooks.

Computers have appeal for students of all ages because it is more fun to practice using a computer. It is a nonthreatening environment where the students can easily check their progress and learn from their mistakes. This is particularly important for students with LD because they often need more practice in order to learn and remember a concept. Sound and speaking the text can be combined with visual images. Record keeping features and flexibility allowing a student to progress at his/her own rate can be built into the program. Educational software for mathematics has unfortunately, up to now, been mostly limited to game type programs or drill programs. Most of the sounds that are used are distracting rather than information giving. When well-designed computer-based and video instruction is provided on a daily basis, children with learning problems can develop declarativ procedural knowledge in basic math skill equivalent to that of their nondisabled peers. (Gold Mertz & Pellegrino, 1989)

Educators have known for years that understanding of mathematical concepts is crucial to long mastery and use of mathematics. The new NCTM (National Council of Teachers of Mathe standards reflect this philosophy. There is a need for *all* students to learn to solve problems, reason learn on their own. (Goldman, Hasselbring, & the Cognition and Technology Group at Vanda 1997). This need for teaching for understanding was one of the crucial ideas behind the developm the MathPlusToolBox.

4 DESCRIPTION OF THE MATHPLUS TOOLBOX FEATURES

$ The MathPlusToolBox uses graphics, sound and captioning so that all students have acc to the concepts being taught. There are a variety of tools to help a student learn and pra basic math skills. The Toolbox will run on a Windows 95 machine with a SAPI com speech engine.

$ The MathPlusToolBox has a universal design. One of the major problems with comm math software is that it is inaccessible to many students. Students with visual impair can't see the screen. Students with hearing impairments miss all the auditory inform Students with learning disabilities often get confused because they need both auditory visual display without a lot of extraneous and distracting material.

$ Audio is used in several ways. The speech engine reads the highlighted text in the scrip in the message boxes, the buttons on the screens, and the problems. It also is used to ex actions being taken in the demonstration boxes. Sounds are also used as reinfor When the answer is correct, the sound is different from when the answer is incorrect.

user selects how much or how little auditory feedback he wants to hear.

$ Captioning allows the hearing impaired to have access to all the spoken text.

$ The screens are relatively uncluttered, and the format is consistent for easy navigation. The graphics are simple and the colors used are consistent throughout the various tools. For example the color green is used for the value of ten's column in the counting tool, and the same color is used for the tens' geometric figures in the demonstration of an addition problem.

$ The type and difficulty of problems can be controlled by the user. Practice problems and flashcards can be controlled via the option's button. A script can be written to generate problems and control the sequence of a student's progress.

$ Tools to aid authors in writing scripts are included as part of the ToolBox. From a pull down menu, teachers can insert script commands into the text. The text is easily accessible making changes possible within a short time frame.

$ Graphics are used to support the development of mathematical concepts. An example would be the use of the graphical demonstration of an addition problem. In a recent journal, the importance of showing the relationship between fractions and decimals was pointed out. To reach higher level understanding, it is important to include graphics and to have the student accustomed to thinking about understanding. (Stein & Smith, 1998) This concept is the basis for the development of the various tools in the MathPlusToolBox.

5 DESCRIPTION OF THE MATHPLUS TOOLBOX COMPONENTS

Counting Tool

- The counting tool looks like a bar graph with four columns divided into ten secti^ Each column represents a number place, ones, tens, hundreds, and thousan[illegible] appropriate number of sections is filled with different colors to match the nu[illegible] written above. A number can be changed by entering a different number in the nu[illegible] box above the graph or by using arrow buttons below the graph to increase or de[illegible] the number of filled spaces and change the number value. The computer will read the entire number upon request.

- It is also possible for a student to see a computerized number line. The number line used to demonstrate a math problem. When the student moves the number line [illegible] the problem and its solution change.

Flashcards

- The flashcards have a design similar to the commercially made cardboard product^

- It is possible to generate flashcards for addition, subtraction, multiplicatit^l division. The type of problem can be changed within the flashcard tool.•

- Problems are generated randomly at the level selected. Two digit problems ca[illegible] generated.

- Students answer in a quick left-to-right manner.

- If a student encounters a problem which is too difficult for them to do in his/her he[illegible] is possible to go directly to one of the work boxes to solve the problem.

- A score box keeps track of the number of problems completed, the number of problems correct, the number of problems missed, and the time elapsed.

- The problem is read for the student. If the answer is incorrect, the student gets the problem again. The correct answer is read.

Math Work Boxes

- There are work boxes for addition, subtraction, multiplication, division.

- These boxes allow a student to do complicated problems on the computer using sound and graphics as an aid for solving the problem.

- The cursor can be selected to move automatically to the correct space. If the answer is incorrect, the cursor doesn't move. Red numbers are used for carrying and borrowing.

- When the student has completed the problem, it can be checked immediately. If the answer is incorrect, the student is returned to the problem. If it is correct, the problem and the answer are read.

- Students have several options from the Work Boxes. A demonstration can be requested. This allows a student to see a working graphical representation of the problem. The computer will work the problem and show graphically how the solution is reached. Another option is for the computer to work the problem for the student. These options could be turned off in a testing situation.

Fraction Tools

- The fraction tools are just now being developed. The tools that are under develo will help the student to better understand fractions in various ways. A simple fra written both in its fraction and decimal form as well as illustrated with graphics.

- Other tools help a student to understand equivalent fractions, factoring, and the add of fractions.

Worksheet Generator

- Random addition, subtraction, multiplication, division, problems can be generated printed.

Calculator

- There is a talking calculator available.

Scripts

- Scripts can be used to teach concepts.

- The MathPlusToolBox components can be integrated into the script. Though the to are designed to help the student learn the concepts, the scripts allow more explanations to be given when a student is learning a new concept.

- The type, level, and amount of problems ean be controlled through the use of scripts

EXPERIENCE WITH USE

Dr. Lundquist has been responsible for the programing, the design, and the implementation of voice in the ToolBox. Mrs. Gardner defined the educational concepts of the ToolBox and has made design suggestions from an educator's viewpoint. Several of her adult basic education students have used various components of the workbook. Changes to the program have been based on the input and discoveries made during its use. Several scripts have been written to help students with specific needs. Since, at this time, the MathPlusToolBox is under development, no official testing has yet been done. All students have found using the MathPlusToolBox to be an interesting way to work on math concepts which previously they found difficult to learn. Using the computer flashcards is less demeaning and more fun than using the old fashion cardboard ones. The graphics and the tools help all students to better understand mathematical concepts. The MathPlusToolBox seems to appeal to all age and ability groups.

FUTURE PLANS

The work on the MathPlusToolBox has just begun. It will become an even more powerful tool as it evolves. More tools for teaching and working with fractions, and tools for teaching decimals, and percents will be developed. Word problems will be included. As the MathPlusToolBox is being used, feedback from students and teachers will be incorporated into the product. There is a need to integrate a Braille display or tactile pad so that demonstrations have more meaning for students who are visually impaired.

ACKNOWLEDGEMENTS

This project was funded in part by the National Science Foundation grant HRD - 9554453 (CG) and in part HRD-9452881 (RL).

9 REFERENCES

Gardner, C. (1996) Assistive technology and learning disabilities, http://www.rit.edu/~easi/easisem/gardnec.html.

Gardner, J. (1997) Overview of the Science Access Project, http://dots.physics.orst.edu/overview.html

Goldman, S.R., Hasselbring, & the Cognition and Technology Group at Vanderbilt (1997) Achieving meaning mathematics literacy for students with learning disbilities. *Journal of Learning Disabilities*, 30-2, 198-208.

Goldman, S. R.., Mertz, D. L., & Pellegrino (1989) Individual differences in extended practice functions and solution strategies for basic addition facts. *Journal of Educational Psychology*, 81, 481-496.

Raman, T.V. (1997) On the internet, no knows you're no a dog either! http://simon.cs.cornell.edu/home/raman

SRI (1997) NSF's study of individuals with disabilities majoring in science, mathematics, engineering, and technology. *Pamphlet*, Menlo Park, CA.

Stein, M. & Smith, M.. (1998) Mathematical tasks as a framework for reflection: from research to practice. *Mathem Teaching in the Middle School*, 3-4, 268-275.

10 BIOGRAPHY

For the past five years, Carolyn Gardner has been researching the uses of technology for students with learning disabilities. She has presented at several conferences and written several papers on this subject. For the past twelve years, she has taught adult basic education at Linn-Benton Community College.

After receiving his PhD in 1993, Randy Lundquist joined the Science Access Project at Oregon State University. He is the author of Triangle, a computer program that permits blind people to read, write and manipulate scientific information.

AGENTS IN AN ADAPTIVE LEARNING ENVIRONMENT FOR SPECIAL NEEDS EDUCATION

Jaakko Kurhila[1] and Erkki Sutinen[1]

Abstract

In a learning situation, varying needs of disabled children can be met in a personal and adaptive learning environment. We give a description of an agent-based learning environment for special needs education. The learning environment ensures individualized learning processes for every learner. We have implemented a functioning prototype targeted to children suffering from frontal lobe lesion. Although comprehensive tests with the prototype are lacking, initial reactions from the special educators have indicated that the concept is appropriate and worth developing. Modular and open design of the prototype provides a basis for further research.

1. Introduction

The concept *learning environment* can be viewed as a physical space where learning occurs. In our framework, the learning environment consists of networked computers with accompanying software and their users. Users can be learners, teachers, special educators, occupational therapists, neuropsychologists etc.

The basis for this framework is to tailor an *agent* for each participant in the learning environment. The definition of an agent is somewhat obscure [2, 3]. In our framework, an agent can be defined as a part of a program that acts on the behalf of the user or, as Russell and Norvig state it [4, p. 31], "an agent is anything that can be viewed as perceiving its environment through sensors and acting upon that environment through effectors."

The framework is to be used in a normal classroom setting with one or more teachers and several learners. The framework is designed to support the learning and teaching processes of each individual learner and teacher. One of the goals is to divide the responsibility of the teaching between a computer and a human teacher.

[1]Department of Computer Science, P.O. Box 26, FIN-00014 University of Helsinki, Finland

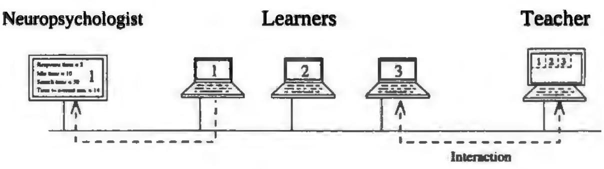


Figure 1: Distributed learning environment, where every participant has his/her own agent

There are customized agents for each learning environment participant. The agents belong to vari-ous classes according to environment users. Every learner has his/her own *learner agent*, teacher has his/her *teacher agent* etc. The agents communicate with other agents and modify the teaching and learning processes. Learner agents can, for example, provide exercises and additional material to learners when needed. Teacher agent observes the learners and provides information from the learners' progress to the teacher. Other agents, e.g. *neuropsychologist agent*, can gather information from the learners' actions to the neuropsychologists and help to diagnose the user.

Figure 1 shows a simple example of a possible learning situation. The teacher agent has found learning difficulties with the learner in location 3, so he/she is instructed directly by the teacher, while learners in locations 1 and 2 are kept activated by their learner agents. Neuropsychologist agent gathers data from the actions of the learner in location 1.

The first implementation of the learning environment is ready. It is a prototype of the learner agent for one learner, and it is to be used for practicing elementary mathematics with individualized adaptive learning path. Because of the open design, the subject can be changed from mathematics to any other desired topic.

As stated in [1], there are no sophisticated computer science solutions to special needs education, but the need for proper software is obvious. Practical reasons include the lack of human teachers. With software agents, not only the quantity of the teaching can be enhanced, but also the quality and depth.

In the following section, we outline the intended users of the prototype learner agent. In chapter 3 we describe the prototype in detail. In the concluding section we give an early evaluation of the learning environment.

Intended users

Special needs education differs radically from ordinary education because of the heterogeneity. For our prototype implementation, we have limited the targeted users to children with *mental programming disorders* [5] caused by frontal lobe lesion, but the schema is most likely to be useful with other types of disabilities even without alterations.

During a learning situation, mental programming disorders appear in various forms. The learners have difficulties in concentrating during a long-lasting task and upholding their motivation. They must be led to the final goal by dividing the problem into smaller and shorter subtasks. The lost motivation can be brought back by personal, emotionally influencing feedback. Activeness can be regained by multi-sensory feedback. Computers are able to convey visual, aural and possibly tactile feedback.

The line between essential and non-essential is vague for frontal lobe lesion patients. They cannot find essential points in a complex problem. Therefore, presentation must be eliminated to the simplest form. Even a slightest misfit distracts or interrupts the execution of the task.

3 Prototype implementation: the learner agent

In this chapter, we describe the preliminary implementation of the learning environment for one computer. The objectives for the prototype and its architectural solutions are presented in detail. Visuality and user interface issues are examined briefly.

3.1 Main objectives

The main objective in the prototype is to practice elementary mathematics, i.e. addition, subtraction, multiplication and division. It is presumed that the learner is already familiar with the presented arithmetical operations, or is taught by a human teacher while using the program. In other words, the learner knows how to add to numbers together when practicing with the addition calculations. The agent offers a practice partner to rehearse mathematical problem solving. If the learner finds trouble during practice session, he/she should be helped by the computer. In this framework, the learner agent is a program coordinating the practice session. The agent decides which exercise is presented to the learner by analysing the learner's behaviour. To make a decision and to fetch an exercise the agent uses external changeable databases.

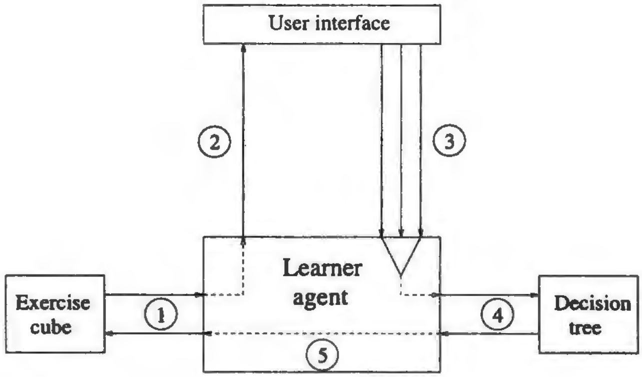


Figure 2: The logical structure of the prototype.

There can be several learners and one teacher, but they do not interact with each other by a com[illegible] The teacher's task is to monitor the learners and their actions, and provide help when needed. In [illegible] future, message passing and other kinds of cooperation can be included in the environment, to ena[illegible] collaborative problem solving.

Multi-sensory feedback is achieved by presenting the feedback visually on the screen and a[illegible] by loudspeaking the exercises and feedback. To avoid distracting other learners in the same ro[illegible] headphones can be used. Tactile feedback is not implemented in the prototype.

The program keeps a score from problem-solving so that the learners can observe their personal p[illegible]gress. Points are given according to the level of difficulty. The points are not stored in a publ[illegible] available high-score list so that a learner can compare him/herself only with his/her previous acco[illegible]plishments.

The prototype is coded with Java, and it is usable in common windowing environments such as [illegible] Windows or UNIX X-Windows.

3.2. Functionality of the learner agent

The logical action in the prototype learner agent goes as follows (Figure 2). When a practice se[illegible] is started, the first exercise is fetched from the database called the *exercise cube* (1) and provid[illegible] the learner agent through the user interface to the learner (2). As the learner uses the software, he/[illegible] generates data which is gathered by the learner agent (3). The data the learner agent gather[illegible] be response times, correctiveness of the answers, number of attempts, and so on. The learner a[illegible]

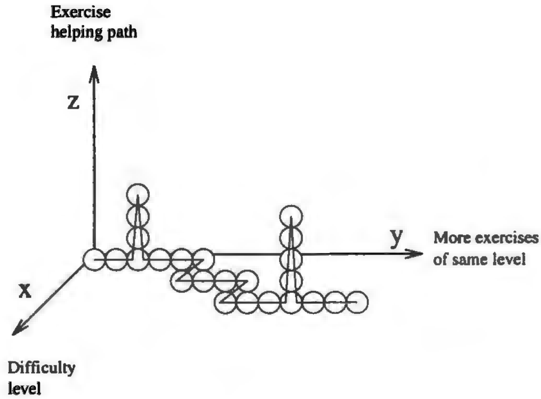


Figure 3: A visualization of a learner's trail through the exercise cube.

ks the data through a *decision tree* to find out how the learner should be helped or which exercise i fetched next from the exercise cube and presented to the learner (4 and 5). This procedure is ated as long as the execution of the program is terminated or the learner agent runs out exercises. tailed descriptions of the exercise cube and the decision tree are given in the following sections.

ogram can store each learner's progress and individual achievements, so that the learner may ontinue from the level he/she has achieved. The learner identification is done with a typical login dure.

3.1 xercise cube

JSie cube is a term for the set of exercises in the prototype implementation. In this prototype implementation, the exercises are mathematical problems. The term cube is used because the set has three erent dimensions. The first dimension is the level of difficulty (x-axis in Figure 3). The proper vel of difficulty is inferred from the amount of terms in the problem, the magnitude of terms, the itude of answers, the amount and closeness of multiple-choice answers, etc.

he second dimension in the cube is the set of exercises of equal difficulty (y-axis in Figure 3). A rge number of exercises and random picking method within that level ensure that the session is not milar to previous sessions, e.g. when the learner has reached his/her level and is continuously trying 6 advance to the next level.

he exercises can often be partitioned to smaller problems or subtasks. For example, if the learner

cannot solve a problem 3+2+4, the problem can be partitioned to 3+2. If the learner can answ
the original problem can then be presented as 5+4. This is the third dimension in the cube a
called the *exercise helping path* (z-axis in Figure 3).

The helping path is not necessarily a partitioning of the problem. It can also be some othe
help the learner to achieve the goal. For example, when the learner is not able to calculate
possible way to help is to present the problem as 5+1.

Every topic has its own exercise cube. The prototype contains cubes for addition, subtract
tiplication and division. Also combinated cubes for addition/subtraction and multiplication/di
exist.

The exercises and their multiple-choice answers must be coded manually to the prog
hand. The reason for this procedure is that problems with a sound pedagogical background are
impossible to create by a computer. In addition, it is not enough that only problems have a sound
agogical basis. Also the multiple-choice answers and the amount of them must be thought thoro
Creating problems manually forces to think the pedagogical basis for each problem and transfe
responsibility to the educationist. However, certain tools to aid the design of the problems to the
can be done in the future.

The contents of the cube can be changed to a different set of exercises, since the cube is an
text file consisting of the actual problem with a proper difficulty ranking, the solving path a
answers. The exercises can include non-numerical terms or even be non-mathematical as long
are multiple choice questions. In the latter case, the prototype contains no support for speech

3.4. Decision tree

The decision tree contains branches of questions which affect the decisions and actions the
agent makes. For example, if the learner has answered the last five times correct and has not consu
too much time, the decision could be to raise the level difficulty. If the learner does not know the
answer, the decision is to present the next step from the exercise helping path. Therefore, each
is likely to go a different trail through the cube.

Figure 4 shows an example of a decision tree. The decisions are based on the answers (arcs)
questions in the branches (nodes). The tree in Figure 4 is a simplified version of the decision

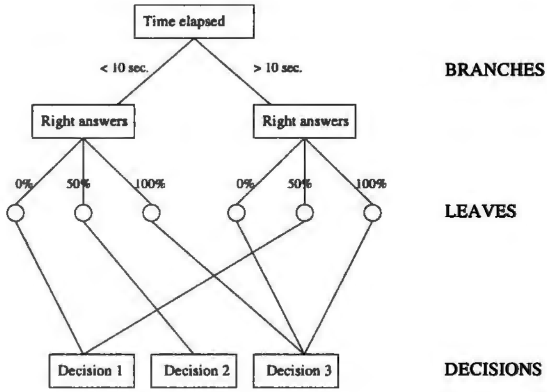


Figure 4: A simplified example of a decision tree used in the prototype implementation.

the prototype. The prototype contains a decision tree with five levels of nodes and 72 leaves which lead to five different decisions. The five possible decisions in the prototype are

- to get an exercise of same difficulty level
- to raise the difficulty level
- to help the learner (i.e. present a next step in the exercise helping path)
- to go down a level
- to send a message to the teacher[2].

The first four decisions steer the learner through the exercise cube. The fifth is used in a situation where helping the learner is beyond the capabilities of the learner agent, or the learner stops responding.

The decision tree can be viewed responsible for the pedagogics in this learning environment. Depending on the questions in the nodes and the limits of the answers in the arcs, the learning environment can be altered to act according to various pedagogically justified learning strategies.

As well as the exercise cube, the decision tree can easily be modified or even totally replaced, because the decision tree is a simple ASCII-text file. The tree in the prototype is static (i.e. constructed before the execution of the program), but a shift to a dynamic tree is one of the future research topics.

[2] Since the *teacher agent* is not yet implemented, this decision is actually useless.

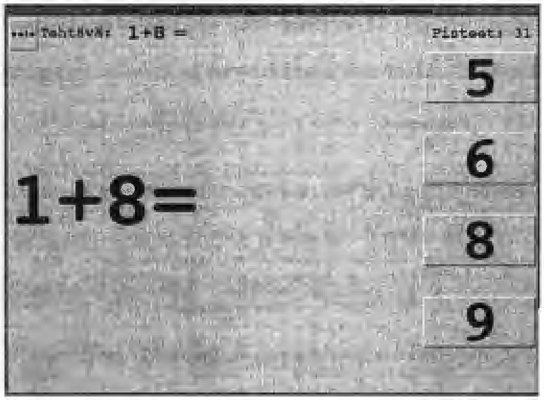


Figure 5: The user interface of the Finnish prototype.

3.5. User interface and visual appearance

The visual appearance of the program is as simple as possible to ensure that the attentive distracted by inappropriate or misleading visual cues and signals (Figure 5). There are no mov on the screen while presenting an exercise, and the colors and fonts are chosen to suit perso slight visual impairments.

The multiple-choice answers are coded with the exercises. The amount of the answers can vary 1 to 10. The answers are always arranged in an ascending order to avoid the mental strain causes to frontal lobe lesion patients.

The user interface is capable to utilize a variety of input devices. The program can be used mouse or some other pointing device. The learner can also control the program from the keyb just two buttons; one to scan the multiple choices and one to select the answer.

4. Concluding remarks

We have designed a framework for an adaptive learning environment to special needs educate prototype implementation for single learner is ready. New features in our approach include the of software agents monitoring and guiding the learner, thus providing the adaptivity needed to meaningful individual learning process. The design has been an interdisciplinary effort.

The prototype offers a practice partner to rehearse basic mathematics, but due to the modu design, other subjects are possible. Also, the prototype can easily be extended to serve as a diagnostics tool for neuropsychologists and special educators. Our approach rooted in special needs education generates new perspectives to "ordinary" education. The designed scenario is sensitive to differences between the learners.

The presented framework is particularly fruitful subject to computer science, since there is an intrinsic correspondence between human problem solving and computer programming methods. This correspondence is even more distinct with the intended users of the prototype, the frontal lobe lesion patients. Even though actual usability and utility tests are lacking, feedback from the special teachers indicate that the concept is genuine and a need for this kind of a learning environment exists.

5 Acknowledgements

The authors wish to thank Tuula Eriksson for her advice on neuropsychological aspects and Erkki Biinranta for his valuable information on special education. We are also grateful to the software project team Sampo Jokinen, Mikko Koskenniemi, Sami Laasanen, Ran Nyman, Tuukka Vartiainen and Pasi Väisänen for the implementation of the prototype.

References

[1] ERIKSSON, T., KURHILA, J. and SUTINEN. E.: "An Agent-Based Framework for Special Needs Education". In SCAI '97 Sixth Scandinavian Conference on Artificial Intelligence: Research Announcements, Helsinki, Finland, August 18-20, 1997, pp. 3-12. University of Helsinki, Department of Computer Science, Report C-1997-49.

[2] FRANKLIN, S. and GRAESSER, A.: Is it an Agent, or just a Program?: A Taxonomy for Autonomous Agents. In Proceedings of the ECAI'96 Workshop on Agent Theories, Architectures, and Languages: Intelligent Agents III, pp. 21-36, Springer, August 12-13 1997.

[3] NWANA, H.S.: Software agents: an overview. The Knowledge Engineering Review 11, 3 (1996), 205-244.

[4] RUSSELL, S. and NORVIG, P.: Artificial Intelligence: a Modern Approach. Prentice Hall, 1995.

[5] VILKKI, J.: Neuropsychology of mental programming: an approach for the evaluation of frontal lobe dysfunction. Applied Neuropsychology 2, 1995, 93–106.

New techniques in education
A status report of the Austrian situation

Prof. Dr. Johann Guenther
Danube-University Krems

We feel the changing needs of the economy but our education system does not react to it. Traditional forms of education are continued. Education systems stay the same.
Education has to adapt to the changes of the labour market, in standard education as well as in a increased postgraduate education.

1.Education and postgraduate education - present level

Education and postgraduate education never ends. The basic education in schools is only a part of lifelong learning process.

1.1.Postgraduate Education - half-life period of education

A company in the high tech field will use 80% new technologies in 5 years. Employees have to be trained in these new technologies as due to fluctuation and retirements only 20% new employees will enter the company. Therefore 80% of the old employees have to be trained for the new challenges in order to secure the competitiveness of the enterprise.
The only point is to hold the present market position with which no new markets or segments are developed.
This makes too long education times senseless. A study which lasts longer than five years imparts knowledge which is already useless after graduation.
To remain competitive postgraduate education has become part of our education system.

Although the number of youths will decrease by 33% between 1985 and the year 2000, the number of employed persons will increase. We count more and more employed persons in middle working age for whom postgraduate education is important and whose knowledge is needed by the economy because there are not enough youths to follow.

1.2.Education organisation

In the international economy flat hierachies have succeded in order to react faster to changes. Between the general managers and the customers of an enterprise there should be few steps in the hierachy. With „Business Process Reengineering" some enterprises even try to eliminate hierarchies so that employees work like independent small enterprises.

Our school system has not yet reacted to that development. Between the „customer" pupil and the Education minister there are five steps: the teacher, the director, the school, the school inspector of the district, the school inspector of the country and the ministry.

1.3. Monopoly of teachers

Our children will be faced with an unstable world and future. Times in which people learned one job and worked in the same profession until their retirements are over. They will change their profession frequently. They will be unemployed very often and the secure employment will be more and more unfamiliar.
How can these people be prepared for their future?
By teachers with civil service status. By teachers who know no competition. Tousands are unemployed and cannot work in their professions. To some extent they are better than their working older colleagues but they do not get a chance to demonstate. The old do not need to improve or make efforts. A Viennese teacher made an inquiry and found out that in reference to information technologies of all teachers between 25 and 45 years - hence those who will still educate for some years - only 15% can operate a text processing program and 7% a calculation program. This is how our children are prepared for the future! How long is it possible to maintain this protecting Monopoly?

1.4. Education in languages

Society is getting more and more international and global and therefore needs better educated people regarding languages. English as the only foreign language is not enough. For middle management functions a second foreign language is necessary to participate actively in the European economy.

1.5. Economy - Market - Professional Life

Education is a product and has to be market orientated. Education systems have to react to changes on the market.

1.6. Shiftings in the potential of manpower

The tendency away from agriculture and forestry to a service industry is growing steadily. With the opening of the Eastern borders a mass of cheap labourers enter the European Union. In order to stay competitive it is necessary to intensify the qualification of our labour force. By means of rationalisation we are not able to compensate these differences in costs. We should try to differentiate in quality and we should concentrate on qualified branches. Holding on to the old economy sectors makes us automatically more incompetitive.

in %		1951	1961	1971	1981	1991	2001
agriculture and forestry		33	23	14	9	6	4
industry and trade	41	47	43	41	36	30	
service industy		26	30	43	51	59	66

working persons according to economy sectors in percent

Until the year 2000 the demand for employees with unqualified and low qualified skills will re
Middle qualified jobs will stagnate but higher qualified jobs will rise significantly.

1.7.Who will be needed and what should be thaught?

The demand for low qualified workers is getting less. Some skilled labour professions - specialists with EDP knowledge - will be needed furtheron and in some skilled labour profess apprentices are required.
A demand exists in economy orientated services like consulting in law and economy, processing, marketing, advertising, research and social orientated services.
High qualified employees are in demand. An actual example is telecommunication. After the the post monopoly this branche is booming and creates a vacuum for skilled telecom workers.
Professional life therefore needs fewer workers in factories and production places but installation and servicing jobs.
Jobs in economic and public services like culture, entertainment, health and law are in demand.

1.8.Social competence

In international statistics Austria ranks high in education (compared to industrial countries of „Western World).
Only 2,3% of all those aged 15 do not decide for further education, two-thirds of them are This education stop will confront them in the future much more with unemployment than boys.
We also lead in the age group of those aged 20:

- have finished their apprenticeship
- have finished technical college
- a professional higher school
- a higher school
- compulsory education

We still do not exploit this international advantage because our education lacks competence". Our pupils and students are confronted with a lot of specialised knowledge but not thaught how to make use of it in practice. Only 20% of those students and pupils have „social competence" to put their knowledge into practice.

This form of education is caused by the missing period of promoterism. It should be our aim everybody becomes independent to achieve self-realization.

This tendency is shown very drastically in the group of university graduates: over 70% of university graduates have become teachers, lawyers and civil servants.

' fcitiians have to regain the courage to become self-employed with a business.

M.Single Fighter

h our complex world of new technologies the economy always needs more group dynamic nployees. Individuals are seldom able to solve a problem on their own. They have to work in ;ups. We receive our last group dynamic education in nursery school. In primary, secondary and Ihgher school as well as academic education single fighters are raised, who are later not needed in Iteconomy.

1.10.Combined Education

Secondary schools try to give pupils a general horizon. A specialisation has to take place later. The jeducation from highschool has to be combined with a special training, an apprenticeship or a niversity study in order to increase one's market value. Therefore 80% of all highschool graduates (o university. Only 40% of all graduates of specialised highschools go to university, the other ÜWget a job in the economy after graduation. Technical colleges are an ideal form of education dich lead to a specialisation after highschools.
Colleges offer a specialised training and may therefore replace apprenticeships.
A single apprenticeship is not sufficient anymore. Not only special occupations like mechanics or ÍBdresser but combinations with other occupations will be needed. Then the single worker will toe more job security.

Information and communication technologies in education

A globalisation of electronic media creates new possibilities which go beyond the traditional forms nlEducation, the reading of books and frontal education. Due to electronic nets connections are lsit up for which there were no possibilities in traditional systems.
IE exploding knowledge of the last years will be linked in a net. Knowledge will be made Wilable and usable. In the past few years technology has helped in the production of knowledge, writ supports in the selection and processing.
to the last 20 years we have been blinded by technological innovations. Today we are re-thinking adare using these technologies. The technical knowledge is becoming part of our lives,
h 1996 25% of all households and 40% of all teenagers in Austria owned a personal computer. 9Wof them have CD Rom or a Modem, so they are prepared for communication, which is also sjiresented in 30 million internet users worldwide.
lie1^talisation of the telephone net is pushed in Europe. On the basis of ISDN a data highway fcvelops. By means of

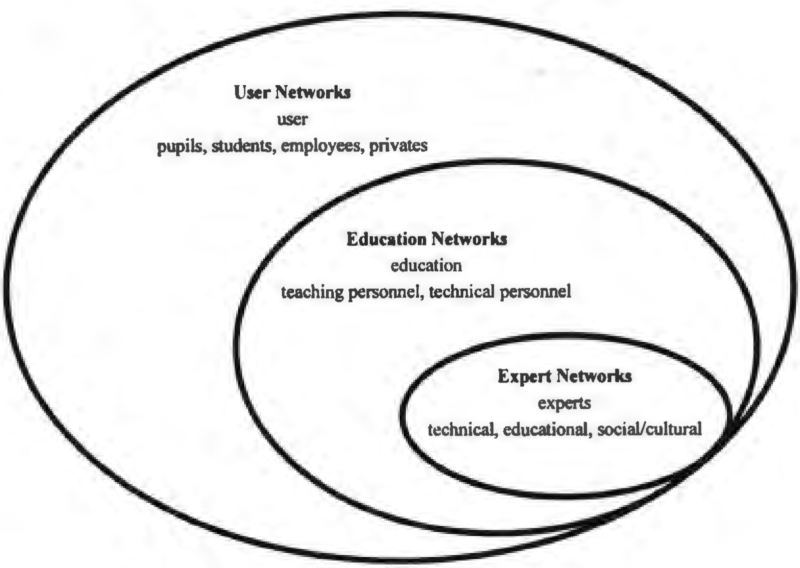
* a digitalised public net
* a digitalised „in-house" transmission and
* corresponding spreading of end units

iebasic requirements are made for a modem education instrument.

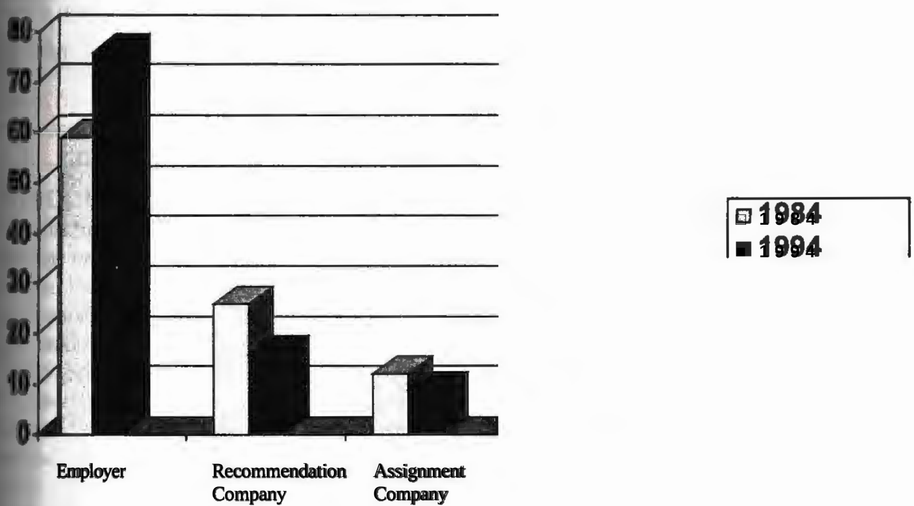
A complete education network should comprise 3 rings:

- The innermost ring should include highschools, universities, colleges and ped academies. In this net specialist and instructors - teachers and professors - will be add
- The middle ring eonnects the places of education. By means of this net information for t is exchanged e.g. common data banks for science, multi-medial education programs administrative software for the school bu
- The most exterior data ring addresses pupils and students. By means of that ring add support to traditional education can be given for learning. Especially the traditional hom can be modernised with the instrument „globalised personal computer“. For a long tim children have been educated to be „single fighters“. In school one learns to succeed individual, to learn an individual and as individuals we are judged, but practice and ec need group dynamic workers, who are able to develop something in a group. The last dynamie education we give our children is in nursery school. The personal compu criticized for making its users isolated, could make homeworks to a group dynamic Over the net the pupil could stay in contact with his classmate and they could solve pr together. He is also able to contact the responsible teaeher to ask for help. Especially in when we register high unemployment of teachers and parents spend billions for private le this „homework net“ could be an up-to-date alternative, whieh is no more expensi conventional private lessons.

3 Rings of the Educational Network



Decision for Postgraduate Education



3. Examples

Internationally there are many examples how the application of new technologies makes teaching more efficient and interesting. Two Austrian examples can be mentioned in this context:

- A primary school in the „Wachau" (Austrian wine-region) exchanges information via Internet with a partner school in China. The children explain to each other their countries: how and where they live, what their parents work, which stores and businesses there are in their town etc. This information is partly exchanged via Internet and partly pupils make a video-film with their teacher in their home town for their friends. With our conventional education in geography it would not be possible to achieve that kind of insight the children get with these new techniques.

- Education at the Danube University Krems: On one day lectures on a particular subject will be held, for example on „telework". An Austrian professor holds a conventional lecture for two hours and explains to the students what telework is, which kinds of telework exist at the moment and what the legal situation in Austria is. Then videoconference connections to different countries will be made. Colleagues from Sweden, South Africa, Singapur, Great Britain and California explain to the students via videoconference the situation of telework in their countries. After an half-hour lecture the students can ask the professor from the foreign country questions for another half-an-hour. After such a day the students leave the Danube University with an international overview about the subject which has been dealt with. An international credibility has been created, which a single local professor would not be able to create even if he had the necessary knowledge.

If we do not provide students with these new achievements, it means that we deny th
resources of the country which is reponsible for education.

Each child is entitled to have a personal computer in school. This demand should be fulfil
next few years.

Playful forms of education will change the access to education. Our children will make pr
from below on the socienty.

A CAI System for Giving Blind Children Spatial Awareness

Ryota SHIMIZU[1], Hiroki MINAGAWA[2] and Noboru OHNISHI[1]*

1 Dept. of Information Eng., Nagoya University 2 Dep.t of Information Science, Tsukuba College of Technology

To help congenitally blind children acquire spatial awareness, we developed a computer-assisted instruction (CAI) system. When a blind user arranges blocks of different shapes according to vocal commands presented by the system, the system analyzes the shape and location of each block by image processing and outputs the result as speech- whether it is correct or not, and the correction, if necessary. By repeating tasks using auditory corrections instead of the help of sighted persons, the mean acquire spatial awareness.

1 INTRODUCTION

Vision substitution systems have been developed to overcome handicaps of the blind in information acquisition, localization, and mobility[1-3]; such as the Optacon, Kurtzweil reading machine, the Eagam reading system[4], mobility aids using the Global Positioning System[5], Sonic guide and a laser cane. These conventional vision substitution systems assume that the users can understand three dimensional (3D) space. However, congenitally blind people can't sufficiently understand 3D space[6]. Therefore, it is important to develop a computer-assisted instruction (CAI) system to provide the blind with spatial awareness. There seems to have been no research conducted in this area yet.

Thus, we developed a CAI system to enable congenitally blind children to understand space; in particular, the ability to perceive location and orientation in two-dimensional space by generating artificial visual feedback from a computer and a camera.

2 CAI SYSTEM

This system enables a blind person to achieve object manipulation in 2D space through a block

Figure 1 CAI system's Overview

arrangement, with the help of auditory feedback. Figure 1 shows the system overview consisting of a video camera, a video capture board, and a personal computer. A simple interface of only two mouse buttons was designed so that young blind children can use this system easily.

We used three kinds of blocks - circle, triangle, and square (the same size of 5cm). Each block was painted a different color; yellow, red, and blue. The system can thus correctly recognize the type and location of each block by using the color information from the blocks.

A blind user is asked to arrange blocks of different shapes according to a problem presented by the system in the form of vocal sounds. When the user completes the arrangement task using the tactile sense, a picture of the block layout is taken by the camera. The system analyzes the shape and the location of each block by image processing, and outputs the result (whether it's correct or not and the correction, if necessary) using vocal sounds within two seconds by improving the recognition algorithm and reducing the image size .

After a discussion with teachers at the blind school, we selected the four tasks which help the blind learn to understand spatial relationships and manipulate objects : *Block Selection Task* , *Block Location Task* which requests the user to place a specified block in a specified position on the table, *Tic-Tac-Toe Game* for increasing motivation toward training[7] and *Hukuwarai* (an old Japanese game) in which blindfolded people

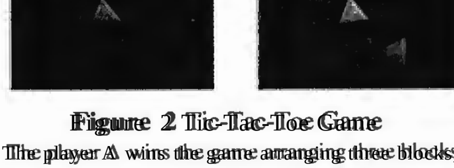


Figure 2 Tic-Tac-Toe Game

The player A wins the game arranging three blocks.

form a human face by arranging facial parts such

*All correspondence should be to Ohnishi: ohnishi@nuie.nagoya-u.ac.jp; Furo-cho, Chikusa-ku, Nagoya Japan

as eyes, a nose, and a mouth. Figures 2 and 3 shows board images in *Tic-Tac-Toe Game* and *Hukuwarai Game*, respectively.

Figure 3 Hukuwarai Game

Mouse □ mislocated in the left is modified in the right.

3. EXPERIMENT & DISCUSSION

We conducted experiments using this CAI system on five subjects. Three subjects were congenitally blind school students (20, 16 and 8 years old), and two subjects who were sighted children (8 years old) with blindfolds.

The results are as follows. *Block Selection Task* was correctly done although the blind child required much more time than the sighted children. As young blind children (about 8 years old) can discriminate simple-shape objects tactically, we must extend the system to enable recognition of more complex objects. In *Block Location Task*, two older congenitally blind subjects and two sighted children could correctly place the block (the rate of 80% to 92%); only the young blind child couldn't correctly place the block (the rate of 57.9%). We believe this is because she has an insufficient ability to perceive spatial images. We also observed two patterns of misplacement with this blind child : a lack of absolute positioning such as "Put the block in the cell first from the left side and second from this side" and a lack of simultaneously remembering all three kinds of information - the kind of block, and the column and row position of the cell.

All subjects enjoyed *Tic-Tac-Toe Game*. While the sighted subjects achieved higher winning rates of 50%(the best of the game), the blind did zero %. The explanation is that congenitally blind children can't understand the blocks' layout and mentally construct an image of it. We believe, however, that congenitally blind children will gain such abilities by repeating these tasks.

Finally, in *Hukuwarai*, four subjects except the oldest blind could form a well-balanced face using the blocks, although the placement time differed among the subjects. The good performance is explained as follows. Because the arrangement area in this task was relatively narrow, they could simultaneously touch four blocks (eyes, nose, and mouth), and easily fo[rm] a mental image of the arranged blocks.

After each experiment, we interviewed ea[ch] subject about the system and the tasks. [The] subjects said that "It was very fun. In partic[ular,] *Tic-Tac-Toe Game* and *Hukuwarai Game* [were] very interesting. I want to play them again". Further, the teachers from the blind sc[hool] commented that tasks involving games are ve[ry] effective in self-learning and training.

4. CONCLUSION

We proposed a CAI system and four tasks [to help] blind children gain spatial awarenes[s. The] experimental results indicated that this sys[tem is] useful for helping blind children learn sp[atial] relationships. The pleasant tasks are partic[ularly] effective. Future research subjects include continuing experiments for congenitally bl[ind] children, investigating whether this CAI sys[tem] is effective in spatial understanding, imp[roving] the system and the tasks, and expanding [the] system to operate in a wider area than p[resently] used.

We would like to thank Mr. Yohichi Hosob[...] and Mr. Akiyoshi Onoda, teachers at Nag[...] Blind School, for their valuable input and [...] with the experiments, and all the subjects [for] their cooperation.

References

[1] H. Ichikawa, H. Ohzu, S. Torii, and T. Wake, "V[...] Handicap and the Support Technology," Nag[...] Press, 1984.

[2] L. Kay, "Electronic aids for a blind perso[n...] interdisciplinary subject," IEE Proc., Vol.[...] pp.559-576, July 1984.

[3] P. Probert, "Report on Rank Prize Fund Medi[...] Technology to Assist the Blind and Vi[...] Impaired," 1996.

[4] Y. Kawai, N. Ohnishi, and N. Sugie, "A Su[...] System for the Blind to Recognize a Di[...] Transactions of the Institute of Electronics [...] Communication Engineers of Japan (D-II), J7[...] 9, 1526-1573(1989-09).

[5] T. A. Herring, "The Global Positioning Sys[tem"] Scientific American, pp.44-50, Feb. 1996.

[6] Flostig, "Flostig's Promote Way of Visual Ab[...] Nihon-Bunka- Kagaku-sya, 1977.

[7] N. Ohnishi, K. Tsuchiya, F. Ito, and S. Nagai, "A[...] of adding factor to the rehabilitation training" [...] Eng., Vol.19, No.5, pp.227-233, 1983.

[8] B. Shneiderman, "Designing the User I[nterface"] Addison-Wesley, 1992, Ch.7.

Session VI

Cognitive Impairments

LEXABILITY FOR LITERACY

Robert Bruce Mahaffey, Ph.D.

Abstract

Lexability is a concept for selecting and using computer software to enhance the literacy skills of persons who struggle with reading, writing, and studying. ***Lexability*** *addresses the pre-literacy needs of young children, the literacy needs associated with all levels of education, and the post-education needs for persons who continue to struggle with reading, writing, and comprehension throughout adulthood. Speech output technologies are key to making software powerful and effective. Appropriate linking of speech with text presents a bi-modal approach of building vocabulary, grammar, punctuation, comprehension, and expression with text-based language. Speech recognition affords the means of expressing thoughts without the constraints of text. Cognitive mapping software provides tools for organizing receptive and expressive thoughts.*

1 Overview

Learning disabilities affect more people than any other disability. Estimates of incidence in the general population vary from 14% to 22%. Reading, writing, and learning problems contribute to most learning disabilities and arise from many different etiologies. Most learning disabilities have begun well before most children begin to read or write and persist throughout adulthood.

Computer-based technologies have significant potential for helping people with learning problems in education, employment and recreation. Speech output technology is an effective tool for breaking through barriers of learning where auditory learning can compensate for textual problems. Speech recognition technology can open gates for people whose expressive language is hindered by writing difficulties. Graphic manipulation of information helps learners with visual perceptual problems and fine eye-motor control problems. Other tools provide drill and practice and general thinking assists.

2 Lexability

2.1 Concept

Lexability is not a product, it is an applied concept that links an array of products to an umbrella strategy for addressing literacy struggles. **Lexability** is a bi-modal approach, based on a theoretical

Oracy to Literacy Continuum, that couples auditory and spoken language with textual reading, writing, and studying processes. The bi-modal approach increases the probability of accurate communication of information between the individual and the computer. **Lexability** assumes that spoken language (oracy) is the foundation for textual language (literacy), and that individuals with functional auditory processes can rely on spoken language to augment the comprehension of text-based language.

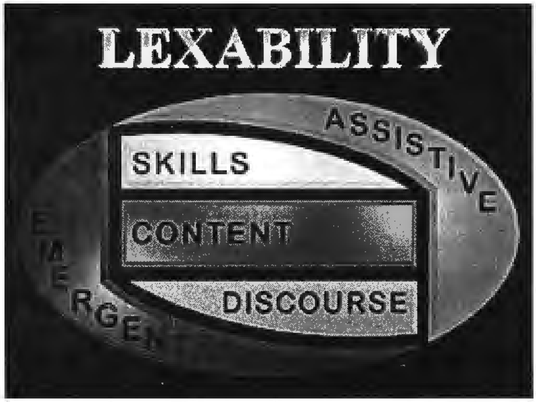


Figure 1. Lexability in a iconic model of Oracy to Literacy

Figure 1, shows an iconic model of the **Lexability** strategy. It depicts a left-to-right continuum that represents birth to death. Each of the five sectors represents a process in overcoming literacy problems. The two vertical division typically coincide with the beginning and the end of formal education. Note that emergent and assistive sectors overlap the formal education period.

The **Lexability** umbrella includes: 1. Therapy-type products that help train individuals to overcome their limitations and maximize their strengths, 2). curriculum content programs that have special bi-modal accommodations to help students with literacy differences, and 3) assistive products that help compensate for differences that cannot be overcome otherwise.

2.2 Sectors

Lexability divides solutions for overcoming literacy struggles into five sectors:

- **Emergent** Practices for developing pre-literacy skills required for listening, speaking, reading and writing.. Some examples of emergent software are: Cognitive Concepts, **Earobics**, that develops phonemic awareness with digitized speech; **IBM SpeechViewer III** that develops speech production skills with digitized speech in and out and with acoustic analysis; Little Planet **Publishing, Thinking**

Out Loud that develops spoken language skills; and Edmark's **Bailey's Book House.** Typically, emergent practices are used in pre-school or the early stages of education, but may continue all the way through formal education.

- **Content** Curriculum content that emphasize the bi-modal approach to learning with speech and text. There are numerous pieces of talking software on the market. It is important that speech and text be synchronized and that the same wording is used in both. Speech provides the auditory enhancement where text is not enough to convey information. Of interest are two specific products, Edmark's **Let's Go Read!** that builds reading and thinking skills through active exploration and **Little Planet Literacy Series** that makes the transition between spoken language and text language.

- **Skills** Exercises for developing basic literacy skills, such as grammar and spelling using the bi-modal approach. Two examples of talking software that develops basic skills are; Skillsbank Corporation's **Cornerstone** teaches reading vocabulary and key concepts in mathematics, language arts or all ages and **Skillsbank 4** that provides remedial coverage or reading, language, mathematics, writing, and information skills for adolescent through adult.

- **Discourse** Discourse activities facilitate the formulation of ideas and then conveying those ideas to others. Software in this area should encourage both divergent and convergent thinking and can include talking encyclopedias, dictionaries, and Thesaurus references. It also includes talking tools searching the web and can also include speech recognition programs such as IBM's **ViaVoice** that allows a user to dictate the content of what is to be expressed and to listen back to what has been recognized. Dictating thoughts may be far more productive than typing for a person who has to struggle with keyboard input. Talking word processors such as Don Johnston, Inc's **Write:Outloud,** coupled with the word prediction software, **CoiWriter.** Both programs use text-to-speech technology to enable auditory processing to augment text processing.

- **Assistive** Assistive technology augments a persons user interface to accomplish more than training or therapy can, alone. The use of assistive technology may start before formal school and often continue through adulthood. Arkenstone's **WYNN** takes electronic text from optical character readers or other sources and makes it more readable both by making the text more visually manageable and by using text-to-speech to read it aloud. It also has study aids that help with comprehension and thought organization. One such aid uses digitized speech or audio annotations of the text. Don Johnston, Inc. has developed the **Discover Series** that allows accessibility, coupled with text to speech that speaks.

3. Summary

The Lexability concept is a construct that categorizes literacy software into five sectors. The purpose of this construct is to help educators visualize, via the Lexability icon, why they are using software and what should be expected of it. Lexability refers to types of software; the software packages cited are good examples of applications for each sector and the examples are not all-inclusive.

ADAPTABLE, ADAPTIVE AND EXTENDABLE INFORMATION SYSTEMS FOR OCCUPATION, ACTIVATION AND STIMULATION OF PEOPLE WITH DEMENTIA

Erwin Riederer [1]
Richard Pieper [2]

Abstract

Dementia is an increasing problem in western countries. This approach tackles the issue if and how Human-Computer Interaction (HCI) science can contribute to improve the quality of life of afflicted persons and caregivers. An entertainment application is developed to enable persons with dementia to spend time alone, a lost capability in this disease. It demonstrates that interaction between persons with dementia and computer is possible, and can contribute to support this target group in daily life. It is argued that only adaptable, adaptive and pro-active system features make such applications possible. The development and implementation of information and communication technologies offers unrealised potentials for home care of elderly with dementia.

Keywords

dementia, home care, information system, entertainment application, adaptivity, adaptibility, extendability, arch model, user model

1. Introduction

This project is conducted in the context of the international EU project "Technology Ethics and Dementia" (TED), a research and development co-operation sponsored by BIOMED2.

1.1 Demography and Prevalence of Dementia

There is an increasing demographic proportion of elderly people in Europe and a rising age of the elderly population. Since the risk of a person to develop dementia is strongly increasing with age, the prevalence of dementia is growing. Dementia is most commonly caused by Alzheimer's disease.

1.2 Technology

Up to now cognitive impaired persons and especially demented persons were neglected as a user group who could profit from computer technology. Beside the fact that the spectrum of tasks done by

[1] Institute for Communications Engineering, University of the Federal Armed Forces Munich, D-85577 Neubiberg., Germany, Tel. +49 89 6004 3919, erwin.riederer@unibw-muenchen.de

[2] Department of Sociology, University of Bamberg, D-96045 Bamberg, Germany

commonly available software scarcely meets the special needs of this target group (utility aspect), its structure produces a complexity of Human-Computer Interaction (HCI) which is not appropriate for cognitive impaired persons (usability aspect). Correspondingly the general opinion is supported that their cognitive impairment makes it impossible for persons with dementia to manage and profit from computer technology (acceptability aspect). The authors argue that the issue of developing applications to improve the quality of life of cognitive impaired persons and their carers is mainly a challenge of HCI science. The goal in this approach is to develop supportive technology for people with dementia and their carers.

2. Methods and Research Settings

The rationale in this approach starts at the focused problem area of dementia, its symptoms, problems and implications. User needs are identified and the specific need addressed in this approach is the need of meaningful activity and occupation to be satisfied by an information system providing stimulating media. The impairments of people with dementia and implications of the social context are transformed into system requirements. An information system in the form of an entertainment programme has been developed which meets these specifications. A prototype of the computer system has been implemented in a diagnostic clinic and day care centre. Initially the patients use the system with assistance by carers in sessions of varying length. The usage reveals functional capabilities and preferences concerning the selection from a broad spectrum of multi-media material. According to this experience an individual user profile is made and used for further system adaptations. Then the patient is animated to use the system alone. Evaluation methods include an automatic, machine generated protocol, participant observation, and interviews after the session.

Although it may seem that the underlying development model is linear and straight forward with sequenced phases of user analysis, investigation of user requirements, system specification, technical development, implementation in social context and evaluation, in fact, it is not, since the phases are interdependent. Without going into details we like to point out that, for instance, the user requirements are already described in terms that are used to specify the goals of the system to be developed. The common background is the problem of technological feasibility to meet these requirements in principle. Also the description of user requirements and the system specification are interdependent - and the most challenging part of finding a possible solution. The impact of a technical system implementation in the context of the 'social triangle of home care' - i.e. patient, informal carer, formal carer - is unpredictable. Apart from the utility and usability aspect the acceptability is crucial. Therefore, the development process must be explorative, use iterative or prototyping methods, and have experiential feedback.

3. Dementia

3.1. Symptoms of Dementia

The symptoms of dementia can be divided into
- cognitive impairments
- emotional, personality and behavioural changes
- physical impairments

The cognitive impairments are:

- memory impairment, especially recent memory, e.g. forgetting information, forget to finish tasks like turning off the stove or the water, forget names, events, occupations, bills or appointments
- difficulty with abstract thinking and reasoning
- loss of judgement
- language difficulty like difficulty speaking or naming objects
- disorientation: Afflicted persons do not know the date and time, where they live. They may end up wandering aimlessly and not find their way home.

Various emotional, personality and behavioural changes can appear:
- agitation, combativeness, depression, fear
- symptoms of psychoses, e.g. delusions, auditory and visual hallucinations, paranoia
- sleep disturbance

Physical impairments are:
- unsteadiness of gait, tendency to fall
- incontinence

3.2 Problems of People with Dementia

The afflicted people have a lot of difficulties in coping with everyday life because of their cognitive impairments. There are various safety problems like injuries, falls, fire risk, wandering and getting lost. A major problem of demented persons is their inability of spending time alone, i.e. keeping themselves occupied. As the majority of them wants to stay at home, the informal carer (in most cases a family member) often is the second victim of this disease. Supervision of the demented person up to 24 hours a day, is a huge burden for the informal carer.

3.3 Central Problem Dimensions

As cognitive impairment precedes physical impairment in the disease the usual sequence of the ageing process is reversed, and cognitive adaptation to new solutions is a central problem. A number of special aspects have to be considered:
- Each afflicted person has its own biography. The target group of people with dementia has a diversity of individual need profiles.
- The time between diagnosis and institutional care is about 3 years. As this 'time window' for intervention is rather small, it is questionable if technological interventions are possible and/or meaningful.
- The habitual, emotional and affective relations to the environment are vital, thus, the situational impact of a technology may pose a decisive problem
- The acceptance by and the support of carers is crucial in the introduction of a technology. The 'target group' is the 'social triangle of home care' (patient-informal carer-formal carer).

3.4 Ethical Dilemmas

As people with dementia have a decreasing competence to reason about and decide on technological interventions someone else has to take responsibility and make decisions. Autonomy, referring to the right of the person to make decisions for themselves, is or may be violated [1]. Competing moral principles (ethical dilemmas) may appear. The approach in [5] proposes a careful weighing of the ethical principles of beneficence, justice and autonomy.

4. Supportive Technology

The objective in this approach is to develop supportive technology for persons with dementia and their carers. The development and implementation of information and communication technologies offers unrealised potentials for the home care of elderly with dementia. There are several advantages of using a computer programme to satisfy special needs (e.g. spending time) in contrast to conventional ways with usual objects and assistive devices or being helped by a carer. In contrast to objects like books software can be adaptive by offering an intelligent active user interface (e.g. context sensitive, spoken hints, active suggestions). Individual adaptations concerning functional and content aspects are possible. In contrast to being helped by a person, a computer system is a tool which is used by the impaired person his/herself. Mastering the use of the tool conveys an experience and feeling of competence. Thus, self-confidence and well-being also may be strengthened, emotional and behavioural problems be attenuated and perhaps even successful coping with other tasks may be supported.

4.1. The Spending Time Problem

This approach focuses on the increasing inability of the patients with the progression of dementia to keep themselves occupied. Principally connected with this issue is the required continuous supervision, but presumably also various emotional and behavioural problems. An entertainment programme has been developed which should be attractive, operable alone by demented persons and providing a substitute for activities or hobbies, which cannot be practised any more. The objectives are:
- Joyful and meaningful activity for the patient
- Stimulation, support and opportunity for interaction and positive feedback
- Relief of the burden of informal care, enhancement of effectiveness of formal care
- Therapeutic, rehabilitative effects, reduction of medical treatment and wandering

Facing the wide spread cognitive problems occuring in dementia, basic usability issues arise. The usual goals of HCI research, i.e. increase of efficiency and productivity, must be broadened and redefined [4]. Not only are productivity and efficiency issues to be managed, but rather the basic possibility of usage - considering the decreasing level of competence - is at stake.

4.2. Key Issues

The key issues are:
- Can demented persons use computers to improve their quality of life?
- What features must such a system have?
- Do the concepts of adaptive, adaptable and pro-active system features enable an interaction between a person with dementia and computer?

4.3. Translating the Needs and Capabilities of the User Group into System Requirements

Multimedia contents should stimulate, activate and keep people with dementia occupied. The information systems provides the presentation of contents and a selection mechanism enabling the user to start a specific one. The easiest human-computer interaction in this context is to let the user choose one option out of a set of options. More powerful search instruments like keyword matching are not appropriate for demented users. The task of choosing a topic out of a offered range can be described by navigating through a menu hypermedia structure which leads to specific contents.

4.3.1 Adaptability

Since people with dementia cannot keep in mind a large range of available options they are not able to operate an unadapted programme which just executes user's commands and operates in a static predefined manner. The diversity of individual need profiles depending on biography demands that the contents of the application have to be adaptable and even extendable. Due to the mental impairments the patients strongly depend on ingrained and habitual memory resources, thus, adaptations to the specific life experience of demented persons, are essential. For instance, a cartoon-like helping agent on the screen - an option provided by the programme - may be appreciated as a funny and useful conversation partner but may also be rejected as childish or disturbing. Natural speech output mostly is very useful but may have confusing impact. Appropriate contents meeting user's interests are at the core of the aim to entertain. Excluding non-preferred topics enables the reduction of the interaction complexity and makes overview easier. The system adaptations reducing complexity by including only the relevant options are assumed to be a condition of an interaction between demented person and computer. Adaptivity in the sense of self-adaptation [3] needs meaningful interaction data and requires successful system use. Therefore a (self-)adaptation of the system to the demented user from the very beginning is not possible. Since the adaptation interaction is a meta-communication about and an abstraction of the task interaction, the demented persons themselves will not be able to perform the adaptations. A different person called configurer has to perform the adaptation interaction. Performing extensions like adding a content (e.g. personal photographs) needs additional competencies. Also programme installation and system configuration have to be managed. To ensure the feasibility of an implementation of a system for people with dementia a usage concept that clarifies responsibilities for all these tasks must be established. Different user roles and tasks appear, and one person may take one or more roles *(Fig. 1)*.

Role	Task
User	performs intended task, e.g. occupy themselves
Configurer	adapts the system to the user
Expert	extends the system according to user's perferences
Administrator	installs application, adapts application to hardware environment

Figure 1: Roles and tasks around the user and their task.

4.3.2 Extendability

Appropriate contents meeting user's interests are at the core of the aim to entertain. To meet user's preferences individual contents may be necessary. Especially including personal photographs has a meaningful and therapeutical background [6]. Considering the requirement of strong individualisation, extending the entertainment application has to be facilitated to enable a practicable usage of this system.

4.3.3 Adaptivity and pro-active Behaviour

According to functional deficits of the user the system must take initiatives and have competence in order to accomplish the objective in the interaction, i.e. to entertain the user. Apart from exploiting knowledge about the user (adaptability), it is necessary to consider the course of interaction to support the user in an appropriate and meaningful way. Adaptive functionality and active help are regarded to be essential in this context. It seems to be promising to use pro-active system features taking the initiative if the user needs help. Adaptivity assessing user's interactions is considered to be necessary to derive useful actions. The system should gradually compensate missing user's activity and competence in a specific situation by providing explanation, advice and suggestions.

Adaptivity refers to the capability of a system to exploit the history of user's interaction and to perform changes in its appearance or behaviour. The objective is to facilitate solving user's task. There are two kinds of changes:

- Long term changes should adapt the system to the user. If the relevant characteristics of the user alter, these adaptations should also alter. As discussed above, implementing this principle in initial system use is problematic with the specific target group, i.e. persons with dementia, because of the incompetence. But if this initial hurdle is taken and the application is used successfully, this kind of adaptivity can perform further helpful adaptations. E.g. the favourite topics can be recognized by consumption time and a start menu can offer the top options to facilitate access to them.
- Temporary changes or actions are considered to be helpful in specific situations which are characterised by the recent course of interaction and the current programme context. These changes will disappear by reaching a new programme context or after a certain time or event, e.g. next session. Conceivable temporary changes or actions are:
  * Presentation changes, e.g. mark used hyperlinks
  * Functional changes like executing commands differently, remove option, offer new options
  * Pro-active interventions like provide hints, explanation, advice and suggestions

5. Scope of Adaptability and Adaptivity

The Arch model for the runtime architecture of an interactive system [7] divides a system into the components of interaction, presentation, dialogue, domain adaptation and domain. Each layer offers service to the lower layer which hides its performance details and decides on its degrees of freedom. Analogously the categories of adapting a system can be divided into the same layers. Adaptability and adaptivity can be realised if each component offers an adaptation interface to facilitate different components to influence these degrees of freedom according to parameters specific to the user profile and the user's history of interaction. A user model seems to be an appropriate entity to construct, store and to supply assumptions about the user to other system components [2].

6. An Entertainment Application

The developed entertainment application demonstrates that interaction between persons with dementia and the computer is possible. The user navigates through a hypermedia tree structure offering topics of entertainment at the higher levels and leading to specific topics like pictures with information and music of e.g. old actors, pictures of sights in cities, paintings, flowers and animals. An option of a hypermedia page can be taken up by direct manipulation via touch screen.

As discussed above, the categories of possible adaptations and adaptive and pro-active behaviour can be divided into the layers of interaction, presentation, dialogue, domain adaptation and domain.

The domain specific component can perform functional adaptations:

- transformation of the hyperstructure, e.g. simplifying the topology by omitting links and restructuring
- adaptation of contents, e.g. omitting parts or switching between different possibilities. The amount and kind of provided information can be varied.

The domain adaptor component can perform active help like explanation, advice, suggestions and temporary modifications like changes of the hyperstructure sensitive to context and course of interaction (e.g. remove unsuccessful links and offer new, very promising links).

The dialogue component can

- use a preferred dialogue form out of a range of equivalent ones
- provide navigation and decision support, e.g. recommend a promising link.

The presentation component shows or hides additional items and switches between different presentation possibilities, e.g. colours, layout, cartoon-like agent visible or not. The interaction toolkit component performs adaptations of input and output like filtering user's input according to individual parameters to avoid misguided inputs (e.g. ignore double-touches caused by trembling while using touch-screen).

6.1. Adaptability

The role of configurer is taken by a relative or carer in a residential home who partly accompanies the demented person (the user) working with the entertainment programme. User's preferences are found out by observation, empathic conversation and as far as the selection of offered contents is concerned by using knowledge about interests, hobbies and experiences of the user. The configurer interacts with the adaptation application to configure an individual user profile.

6.2. Extendability

The person who is responsible to extend the system according to user's preferences is termed the expert (*Fig. 1*). The expert asks the configurer about new contents to be included, searches for suitable material, and prepares the media for implementation. Adding a new topic to the information system should be kept as simple as possible, e.g. creating a directory which includes the picture sequence in a defined format and labels systematically optional and additional text-files with titles and information and optional sounds according to the pictures. The hypermedia concept enables extensions by adding links and content documents. An extension application could facilitate the implementation of the new topic in the hyperstructure. Apart from operating this application, the expert has to use different programmes to edit the documents.

6.3. Adaptivity and Pro-Activity

Recognizing the favourite topics by evaluating consumption time is used to offer the top options in a start menu and to facilitate access to them (long term changes). The users are supported with active help if they seem to be inactive or confused. Context, individual profile and history of interaction are taken into consideration. A simple reactive temporary presentation change which considers only t last input is marking a used hyperlink. To avoid circle navigation, options used twice are temp removed. On user's inactivity, firstly, a hint how to take up an option is performed, then unsuc links are removed and new, promising links are offered and suggested for selection. The promising option is derived according to user's preferences and the already consumed topics durin current session (which are rejected as promising suggestions). When a picture series is selected twic an advice to choose different topics is performed.

7. Results and Consequences

Persons with dementia can operate adapted computer applications up to a certain degree of de Then they need assistance. In a number of case studies with patients in middle and even prog states of the disease the patients showed surprising effects like activation and strong interest in con to usual passive behaviour. Interaction, presentation, functionality and content must be individ adapted according to personal preferences. Applications for persons with dementia are possible meaningful. This target group can profit from HCI concepts fundamentally because only adaptable adaptive interface and system features enable an interaction between demented person and compu The next step will be to extend the evaluation of the entertainment application to a larger user gro to gain more empirical data. Further applications like day and week planner with reminder fun could support the memory and help the demented person planning and co-ordinating activities. application supporting a psychological therapy called 'Self Maintenance Therapy' [6] is in prog Further investigations should clarify what concepts of HCI, e.g. kinds of models and archite proposals, are useful in this context. The domain of rehabilitation engineering should use H concepts, and, on the other hand, HCI research should recognize rehabilitation engineering as excellent application area to tackle new problems and to develop concepts and goals to support hu activities apart from the usual narrow objectives of increasing productivity and efficiency.

References

[1] Downs M., Assessing ethical issues in the use of technology for people with dementia, in Ethical issues in the technology for dementia home care, the Akon Series 'Aging in the contemporary society' Volume 13, 1997.

[2] Kobsa A., User modeling and User-adapted Interaction, Tutorial at CHI'94, Conference on Human Factors in C Systems, ACM press, 1994.

[3] Malinowski U., Kühme T., Dieterich H., Schneider-Hufschmidt M., A taxonomy of adaptive user interfaces in Monk, D Diaper and M D Harrison, HCI'92, People and Computers VII, Cambridge University Press, 1992.

[4] Muller M., Wharton C., McIver W., Laux L., Toward an HCI Research and Practice Agenda Based on Human Ne Social Responsibility, CHI'97, 1997.

[5] Rauhala-Hayes M., Ethics of care work, in Ethical issues in the use of technology for dementia home care, the Ak 'Aging in the contemporary society' Volume 13, 1997.

[6] Romero B., Riederer E., Self Maintenance Therapy (SMT) in Alzheimer's Disease, 12th International Conf Alzheimer's Disease International, Jerusalem, Israel, 1996.
available at http://speedy.et.unibw-muenchen.de/forsch/bt/publ/israel96/israel96.html

[7] UIMS Tool Developers Workshop 92, The UIMS Tool Developers Workshop. A metamodel for the runtime archi of an interactive system. ACM SIGCHI Bulletin, 24(1):32-37, January 1992.

Theory-Based Software Use in Language Intervention

Mary Sweig Wilson, Ph.D. and Jeffrey P. Pascoe, Ph.D.

Abstract

Delays in language acquisition are the most prevalent developmental problem seen in preschool children. The acquisition of verbs poses a special challenge for children with language problems, perhaps because verbs are richly endowed with syntactic and semantic information and play a central role in the development of grammatical competency. We are developing and field testing software designed to facilitate the development of a rich verb lexicon in children at risk for chronic language disorders. The curriculum is based on Chomsky's Minimalist Program and other linguistic research ascribing a central role for verbs in language acquisition. The software features engaging graphics and animation, interface options to accommodate children with special needs, digital speech output, and an intelligent computer-aided training system that uses artificial intelligence to generate individualized intervention strategies. The completed system will include modules designed to increase the size and diversity of the verb lexicon, increase semantic and syntactic accuracy, and promote the proper use of inflected verb forms. We envision the system being useful in clinical, school, and preschool settings, as well as in the home.

Introduction

Clinically significant delays in the acquisition of language are the most prevalent developmental problem seen in preschool children. Using standardized assessment tools with the goal of identifying children at risk for persistent language problems, various investigators have concluded that language acquisition delays affect between 9% and 16% of all two year olds [46]. About half of these children eventually turn out to be "late bloomers" who go on to develop normal language skills by three years of age. The remainder, however, continue to exhibit significant delays in the acquisition of language and are known to be at risk for chronic language impairments and their sequelae [31,42,47]. A recent epidemiological study of language impairment among more than 7,000 kindergartners (5-6 years old) estimated an overall prevalence of 7.4% [61].

Because language delays often appear to diminish spontaneously and intervention resources are scarce, the value of early remediation in language delay has sometimes been questioned. Yet proponents of early remediation argue that intervention is crucial since children do not simply "outgrow" preschool language delay, and the consequences of this handicap are often cumulative. At three years of age, many late talkers manifest social skill deficits that remain apparent at follow-up a year later [43]. At age four, children with a history of language delay are clearly impaired on measures that are predictive of school success [41]. A recent review by Kelly [25] noted that, while many children with language delay appear to move into normal limits of variation in language ability, few actually develop language abilities comparable to peers who have had no previous history of language impairment. These children therefore remain at risk for chronic language impairments, learning disabilities, academic failure, and behavioral disorders [1,4,47,48,55]. Considered together, these observations clearly suggest that the timely implementation of an effective strategy for facilitating language learning and remediating language impairments could have beneficial and far-reaching consequences.

Chomsky's Minimalist Program &Its Implications for Early Language Intervention

When considering the development of software for language intervention, fundamental questions arise: what children must learn when they acquire a language, and how children manage this learning task. To address such questions, we turned to the field of linguistics. Two fundamental goals of linguistic research are to provide (a) an adequate *description* of what constitutes knowledge of a language, and (b) an adequate *explanation* of how humans come to achieve command of that knowledge. Earlier linguistic theories typically focused on descriptive goals and were associated with elaborate and often language-specific systems of rules. A problem arose, however, when it became clear to many that any system of rules of descriptive merit was well beyond the scope of what a normal child could grasp intellectually or learn simply on the basis of environmental input [2,8]. These considerations led Chomsky to develop several iterations of his theory of generative grammar [5,6,7], incorporating the notion that knowledge of language must exist, at least in part, in advance of experience. In other words, (a) children must possess "innate knowledge" about the form of natural language, and thus, (b) language acquisition is greatly simplified because innately defined constraints place strict limitations on the form that any natural human language may take.

A recent milestone in this theoretical evolution is the Principles and Parameters Model. This model suggests that the acquisition of language involves the maturation and use of biologically determined "universal principles" that govern the structure of natural language, together with the setting of "parameters" - a set of options that can account for most of the syntactic differences among human languages [2,7,11,32,45,54]. This conceptual model has become dominant in the field of linguistics, and provides a useful framework for language acquisition research, in part because the model has many heuristic implications. If, for example, the acquisition of syntactic competence is largely a matter of learning the values of a small set of grammatical parameters that distinguish one's native language, then it follows that understanding of these parameters could lead to intervention strategies that are more effective because they specify linguistic experiences that may optimize or correct the process of language acquisition at a fundamental (versus symptomatic) level. In other words, the model suggests that language intervention should emphasize linguistic data that are likely to interact with innate factors that shape language acquisition, and are likely to "set" the grammatical parameters of the child's native language [2,24,35,54].

The *Minimalist Program* is the most recent effort in this long-standing linguistic research program. Its goal is a further reduction in the complexity of linguistic theory, perhaps to the extent that a unifying explanatory theory of language can be proposed. A key feature of the Minimalist Program is a strong "lexical approach." That is, the program not only specifies that each lexical entry must include phonological and semantic properties, but also requires that lexical entries contain syntactic features such as category membership and inflectional behavior, e.g., marking for number, person, and gender. Indeed, the Minimalist Program assumes that words emerge from the lexicon and enter the computational system in a fully inflected state, thus eliminating the necessity of "affixing" inflectional morphemes during subsequent transformational operations [14]. Such an assumption is intriguing in light of the observation that children with specific language impairment have particular difficulty with the acquisition of grammatical morphemes that mark tense and agreement [30,53], and clearly has implications bearing on the design of new language intervention strategies. In particular, this research suggests that language intervention designed to increase the size of the lexicon certainly ought to place considerable emphasis on illustrating the morphological variety associated with each lexical item in various syntactic settings. The design of our new prototype software embraces this notion.

The Challenges of Verb Acquisition

Some aspects of language are especially challenging for children in the early stages of language acquisition, and particularly so for children with language impairments. The challenge does not seem to be associated

T

#' the single word stage in which simple nouns predominate. Most young children grasp the notion that nris can ^bolize objects and people (i.e., referential nouns), and come to use these words without ttulty. With the emergence of verbs, however, the time course and preponderance of errors during verb ifBsition suggest that even the most competent language learners are challenged by the process. Children tad to produce nouns long before verbs, acquire new verbs more slowly than other lexical categories, and rdce only a fraction of the verbs that they apparently comprehend [17, 19]. Comparable findings are diiBdteider conditions that control for word frequency and phonology [33]. Frequency of exposure to verbs taittle currency as an explanatory variable in any event since verbs outnumber nouns three-to-one among the mlfrequently spoken words in adult language [27]. Nor does the sluggish acquisition of verbs seem to be an tafct of some unique property of English. Cross-linguistic analyses of a variety of languages that feature ifcent canonical word orders and degrees of morphological complexity confirm that children simply learn before verbs, have fewer verbs in their early vocabularies, and make more errors with verbs [17,21].

fawns offered to explain why children acquire verbs more slowly derive largely from four basic features fa generally distinguish verbs in the lexicon. The first of these is that verb acquisition often requires dren to associate lexical information with one of many transient phenomena that occur within a dynamic ay of activity. In comparison, noun acquisition typically takes place in ostensive contexts where the child (perceive the referent continuously while a label is provided [60], and appears to be guided by cognitive tass which lead children to assume that labels are mutually exclusive, refer to whole objects rather than |tas and generalize to other members of the same basic-level category [37]. In the case of verbs, however, iis less clear how a child might be equipped to distinguish, in an ongoing sequence of events, that to which (particular verb might refer. When viewing the transfer of a substance from container to glass, for example, fachild may be faced with the challenge of determining which transient action to associate with *hold, pour, ft spill, Œsplash*, but has no conspicuous means of doing so [20]. There is, of course, good evidence that dWen come to exploit relationships between verb meaning, thematic roles, and syntactic structure to infer possible meaning of new verbs [39,62], but this strategy may not be available in the early stages of tapage acquisition and prior to the development of adequate syntactic competence.

.tecond feature that distinguishes verbs and may confound children's efforts to acquire them involves the itinsic semantic complexity of verbs. Semantic information (e.g., implications of causation, change of ■, manner or result, direction or location of action, etc.) is often encoded within and contributes to the ■ning of a particular verb. Verbs that indicate a change of possession, for example, include *give, lend, É, forfeit, & surrender*, each of which is endowed with implications bearing on the circumstances of the ■ange. On hearing any of these verbs for the first time it seems unlikely that their semantic implications pÉlbe^prehended fully. Indeed, that this is the case is suggested by the observation that children initially iBe only a few representative verbs in a particular semantic category [44], and common verb errors IlHitly are attributable to incomplete semantic marking and the overly broad use of a few 'all-purpose' ■is [50]. An additional challenge to learning verbs, then, would seem to involve the refinement of ■ntic representations and the acquisition of semantically more precise terms.

1

lhid feature of verbs that may contribute to the challenge of acquisition involves the fact that verbs are ...ily associated with thematic roles (e.g., agent, patient) and syntactic functions (e.g., subject, object) In learning the verb *kiss*, for example, a child would have to know how to construct a phrase containing te^quired number and arrangement of arguments (the "argument structure" of the verb), be aware of the toic roles required by the semantic features of the verb [13,44], and be able to assign the verb's thematic ties to the argument structure of the verb. These accomplishments are, of course, beyond the abilities of «y young children who must instead rely on their audience to intuit, for example, whether the verb *kiss* is ing used to describe an event or to initiate one. These circumstances also may account for the fact that children's early verbs tend to be semantically simple; e.g., verbs that label an action without referen result, or a state of change without reference to antecedent events [3,23]. Expanding the verb lexicon bey these "basic-level" verbs places considerable demands on the child, who must now acquire not j additional lexical items, but also a fair amount of additional syntactic and semantic information.

On the other hand, during normal development the initial challenge of acquiring the argument structur thematic roles associated with early verbs produces a body of knowledge that becomes a valuable asset verb learning. With experience, recurrent regularities among the interrelationships between verbs, syn and semantics can be exploited by children in order to more accurately guess at the possible meanings novel verbs [18,34,39,44]. "Canonical linking rules" describe the ways in which various thematic roles a syntactic functions can be linked [44,62]. The links are bi-directional such that children who know semantic meaning of a verb can use that information to construct an appropriate syntactic fram ("forward linking," from semantics to syntax), and those who hear a novel verb in a familiar synta context can use that context to deduce a great deal about the meaning of the verb (reverse linking, fr syntax to semantics). Such linking is facilitated further by the fact that verbs which share certain aspect meaning also tend to have common syntactic features [20,34]. The development and use of cano linking rules may account for, at least in part, the acceleration in verb learning which coincides with emergence of early word combinations, and an improved ability to develop tentative semantic interpret of words with minimal input or assistance [12,52].

Finally, a fourth feature of verbs that may present a special challenge for children derives fro complexity of verb inflection (INFL) [17,49,53]. Not only must children learn a verb's argument structu and the thematic roles it assigns, they must also learn how to mark the verb for tense, person, and numb The grammatical morphemes associated with INFL seem especially problematic for children with langu impairments [15,36]. Additionally, the assignment of case and its marking in the English pronominal sy causes problems for these children. Since the Subjective case is assigned by INFL and Objective case assigned by a verb to its Complements, verb INFL plays a critical role in these areas as well.

Verb Acquisition by Children With Language Impairments

The features that generally distinguish verbs in the lexicon seem more than adequate to account for special challenges associated with verb acquisition, and cause it to seem all the more remarkable children so readily acquire a functional verb vocabulary during their preschool years. But for a signi number of children verb acquisition seems to pose a formidable challenge. Recent research has examined some detail the acquisition of verbs by children with specific language impairments (SLI). Amo observations are the following: *First*, children with SLI appear very limited in their ability to acquire verbs under conditions that reliably produce new verb learning in other children [51,52,65]. *Second*, chil with SLI typically use more uninflected verb forms and have tremendous difficulty with verb morph This is perhaps not surprising given that these children also have difficulties with subject case mark word order, and word classes [9,21,36,49,53,64], which suggests inadequate comprehension of the sy and semantic features that are fundamental to the derivation of verb inflection. *Third*, children with SLI to rely heavily on a relatively restricted set of verbs, most of which encode only fundamental sema distinctions and thus may be used to convey a broad range of meanings [50,66]. Not surprisi reliance is associated with frequent errors involving the substitution of verbs that are sema inappropriate. *Fourth*, children with SLI are deficient in reverse linking (deriving meaning from sy When exposed to a novel verb within a sentence and asked to make up or guess a plausible meaning for verb and act it out with puppets (e.g., *"The girl voozes the boy"*), children with SLI did poorly. The act out an event that accommodated the argument structure of the verb, but the assignment of the was inappropriate [62]. This finding is remarkable in view of the presumably important and pervasive wse linking in deducing the meaning of new verbs [18,20,39,44]. An inability to exploit reverse linking fodaeously would certainly impair language development, and from a clinical standpoint this appears to important deficit to target for remediation.

Facilitating Verb Acquisition in Children With Language Impairments

Is it possible to facilitate development of a productive verb vocabulary in children at risk for chronic language deficits? Exploring this issue can be challenging since there are many strategies for the remediation of language impairments, and these are derived from an assortment of perspectives on language acquisition. Moreover, there are few empirical data to indicate the comparative efficacy of various language intervention strategies [10,16,28]. Even so, a notable trend in language intervention in the past two decades has been a trend toward emphasizing social communication skills in natural settings (for review see [10]). This shift stemmed in part from evidence that the need for social interaction can nurture the development of communicative competence [16,48], and embraces the notion that effective language intervention should be based in naturalistic conversational exchanges between the language delayed child and partners who model appropriate language forms and functions [57,58]. We agree that functional communication should be an important intervention goal. It is also important to recognize, however, that children with language impairments almost by definition *have an impaired ability to acquire language during the course of naturalistic conversational exchanges*, and clinical data show that the majority of these children are in need of formal syntactic remediation. Indeed, remedial procedures empirically shown to facilitate the acquisition of language structures (as opposed to the functional use of already acquired language) feature intensive didactic training [10,63,68]. Lexical training seems best accomplished with focused ostensive naming tasks and frequent repetition [26,49,60]. Thus, while activities that emphasize narration and pragmatics clearly have an important role in language intervention, there remains an additional and important role for activities that emphasize the formal aspects of language. Watkins [64] concluded that "...diverse and inclusive approaches to language intervention/clinical practice offer the greatest probability of success..." and she recommended placing special emphasis on grammatical skills, timing interventions in relation to emerging skills, and using verbs as "the centerpiece of grammatical intervention."

A Role for Microcomputer-Based Language Intervention

We believe that the challenges outlined above provide a compelling example of those in which computer technology can be brought to bear and make an important contribution. In addition to obvious benefits, such as using computer animation to represent verb actions [22], in our experience computer-based language instruction programs are ideally suited to provide the structured interactions needed to illustrate the formal aspects of language, and provide a highly cost-efficient delivery system for effective individualized language intervention. Moreover, research has demonstrated the fundamental efficacy of microcomputer-based language intervention strategies. Children with special needs in the early stages of language acquisition can make language gains comparable to those seen during individual language therapy with a speech-language pathologist, using language intervention software with non-professional assistance. Significantly greater language development and improved communications skills have been observed when regular use of language intervention software was added to the ongoing curriculum of a special education classroom [.'09,69,70]. There also is a broad consensus that language intervention, regardless of specific procedures, is more effective when the intervention is engaging, appropriately challenging, requires active participation, and provides a legitimate reason for engaging in communicative behavior [38,63]. Those who have had occasion to witness children working with properly designed educational software generally agree that computer-based instructional programs are well suited to meet these criteria.

Recognition of the potential of microcomputer-based language intervention strategies inspired us to develop language intervention software that is now recognized as a valuable supplement to traditional intervention

strategies used with children. Our software can be used with considerable independence by very
children, provides a stimulating context for learning, and provides a focal point for highly
conversational exchanges. There are other considerations as well. The intensive instruction
children with language disorders often exceeds the amount of time and energy that parents, teachers,
language pathologists, and special educators can devote to this effort, especially since professional
are at a premium under even the best of circumstances. Moreover, the problem is
economically disadvantaged areas where the prevalence of language delay is so much greater. The
Speech-Language-Hearing Association has called the current shortage of speech-language
the United States "a recipe for disaster" (Wall Street Journal, 11/14/95, p.1). From this perspective,
language intervention software can be regarded as a much needed and cost-effective means to
the efforts of clinicians and educators [69,70].

In developing our language intervention software, we also felt it was important to consider that the quality
any computer-based language intervention strategy currently is based on the ability of the
carefully track the performance of each student, select and properly utilize programs that teach
content, and gauge the timing of intervention emphasis in relation to emerging skills. Further,
decisions must be based entirely on clinical judgments and perceptions of what will challenge but
frustrate each child. Such evaluations can be difficult, and especially so for those who have limited
in clinical research or linguistic theory, yet they are critical since even the most effective software can
expected to accomplish little if not used in an appropriate manner. To address these concerns we turned
examination of computer assisted training paradigms.

The Intelligent Computer Aided Training System

The considerations outlined above have led us to include as an additional component of our
effort the integration of an Intelligent Computer Assisted Training (ICAT) system. This system will
possible for the software to be controlled by "artificial intelligence" methodology designed to
efforts of parents and non-professionals who may implement language intervention plans, and
reduce the burden placed on those who may lack clinical or technological expertise. Laureate's ICAT
was inspired by systems developed by the Software Technology Branch of National Aeronautics and
Administration [40] and formally transferred to Laureate under a Space Act Technology
Agreement. We have now developed and field tested several of our own prototype ICAT
software packages designed for early language intervention. Our current prototype selects and
activities that emphasize particular verbs or verb argument structures, and arranges the presentation
material in such a way that the focus of treatment and levels of instructional support are
modified on the basis of a child's performance and capabilities. The system also maintains records
program use and user performance. In thinking about the ICAT system, an appropriate analogy is a
session. Typically, at the start of a session the speech language pathologist will know what material
ought to work on, and how much instructional support the child may require to succeed. As the
proceeds, the training material and degree of instructional support are adjusted continually based on
child's performance. This is essentially what our ICAT system is designed to do. In broad terms, the
system uses three sets of data to accomplish this: optional *Program Variables*, the *Student History*
treatment *Rule Set*.

Program Variables include all programs that are integrated into the system, all activities within
program, and all instructional variables that might be manipulated during a single trial in a
activity. Each activity, for example, may have options governing the number of visual stimuli that
the screen, the type of task that is to be completed, the content of verbal prompts, and whether (a) the
will begin with instruction, (b) the correct answer will be cued, (c) an incorrect response will

...tional feedback and/or a second try, (d) feedback will be provided, (e) reinforcement will occur, and so on. Activities will be programmed so that any of these trial variables can be adjusted at the start of each trial.

The *Student History* is a user database that includes a record of activities that have been used, the Program variable settings that were used, and all results obtained, on a trial by trial basis. The format of these raw data is such that the Student History can be "queried" for information ranging from the number of correct responses in an activity with a particular level of instructional support, to the pattern of changes in average response time over the course of a session. In addition to providing a basis for decisions by the ICAT system, the Student History can be "queried" using a *Report Generator* designed to extract and present meaningful performance summaries.

The intervention *Rule Set* contains the "artificial intelligence" of the ICAT system. The "rules" are meant to be comparable to the strategies that a clinician would employ during language intervention, e.g., keeping the child on appropriate material and providing only as much instructional support as is necessary to insure a sufficient degree of success. The rules generate assumptions based on performance in particular circumstances, e.g., a series of successful trials may signal mastery of a concept and trigger introduction of another; a sudden increase in response time or decrement in performance may signal fatigue. Three notable features of the Rule Set include the following: (a) It is contained in program resources so that it may easily be modified. This greatly simplifies the process of elaborating and fine-tuning the Rule Set during field testing; (b) The Rule Set is not simply a linear "script" or flow diagram. Rather, curricular material is associated with "weights," indices of a child's performance are associated with "weighting adjustments," and various weighting adjustments can be "broadcast" throughout the system in such a way as to exert a measured influence over a multitude of program decisions. Good performance on a particular problem set, for example, may cause weights to be adjusted so that similar problem sets are introduced with less instructional support. The weightings can also be adjusted so as to exert a global influence on program effectiveness. It is possible, for example, to increase or decrease the overall difficulty of the entire program using a simple slide bar that adjusts all weights in a predefined manner; (c) The Rule Set can govern all aspects of program performance, but there are options to partially or completely override the ICAT system.

During ongoing field testing, the behavior of the ICAT system in our prototype language intervention software continues to exceed our expectations. Higher functioning children are quickly able to demonstrate their verb knowledge by accurately identifying verb animations with minimal instructional support, yet at the same time they receive extra training with added instructional support for those few verbs with which they are struggling. The less dependable performance of children functioning at lower levels triggers an increase in the level of instructional support when they respond incorrectly, until they too reach an appropriate training level. Even after a single session there are clear differences in the material being presented to individual children, and the material is presented differently and appropriately for children with different abilities.

When our prototype development and testing is complete, we plan to make available a software package consisting of a series of modules that exploit the computer's capacity for combining interactive multimedia with best clinical practices and pedagogy. These will include a variety of training and testing activities, and will address multiple aspects of verb acquisition and language development. Modules will be integrated into a shell featuring engaging graphics, animated guidance and reinforcement sequences, digital speech output, and user interface options to accommodate very young learners and children with special physical needs. We expect the ICAT component to allow the software to emulate the intervention strategies of a professional speech-language pathologist on several levels: (1) the initiation of interventions aimed at logically defined classes of appropriate material, (2) continuous monitoring of the user's progress on each type of material,

(3) continuous revision of intervention strategies based on performance, and (4) comprehensive di collection. Given the current paucity of controlled research on the efficacy of microcomputer-b language intervention strategies [56], it seems likely that the data collection capability will be a valu component of the system as well.

Acknowledgement: This work is being supported by SBIR grants #R44HD35255-02 and #R44DC02601-02 from the National Institutes of Health.

Literature Cited

[1] Aram, D. & Hall, N. (1989) Longitudinal follow-up of children with preschool communication disorde treatment implications. School Psychology Review, 18, 487-501.

[2] Atkinson, M. (1992) Children's Syntax: an Introduction to Principles and Parameters Theory. Oxf Blackwell.

[3] Behrend, D.A. (1990) The development of verb concepts: children's use of verbs to label familiar and novel events. Child Development, 61, 681-696.

[4] Catts, H.W. (1993) The relationship between speech-language impairments and reading disabilities. Journ of Speech and Hearing Research, 36, 948-958.

[5] Chomsky, N. (1957) Syntactic Structures. The Hague: Mouton.

[6] Chomsky, N. (1965) Aspects of the Theory of Syntax. Cambridge, Mass: MIT Press.

[7] Chomsky, N. (1981) Lectures on Government and Binding. Dordrecht: Foris.

[8] Chomsky, N. (1995) The Minimalist Program. Cambridge, MA: MIT Press, 420 pp.

[9] Clahsen, H. (1989) The grammatical characterization of developmental dysphasia. Linguistics, 27, 897-9

[10] Cole, K.N (1995) Curriculum models & language facilitation in preschool years. In M. Fey, J. Windsor S. Warren (Eds.), Language Intervention: Preschool through the Elementary Years. Baltimore: Brookes, 39-60.

[11] Crain, S.(1991) Language acquisition in the absence of experience. Behavioral & Brain Sciences, 14, 5 650.

[12] Dollaghan, C. (1987) Fast mapping in normal and language-impaired children. Journal of Sp Hearing Disorders, 52, 218-222.

[13] Dowty, D. (1991) Thematic proto-roles and argument structure. Language, 67, 547-619.

[14] Epstein, S.D., Thráinsson, H., & Zwart, C.J-W. (1996) Introduction. In: W. Abraham, S. Epstein, Thráinsson, & C.J-W Zwart (Eds.), Minimal Ideas: Syntactic Studies in the Minimalist Framework. Amsterd Benjamins, 1-66.

[15] Eyer, J.A. & Leonard, L.B. (1995) Functional categories and specific language impairment: a case study. Language Acquisition, 4, 177-203.

[16] Friedman, P. & Friedman, K. (1980) Accounting for individual differences when comparing the effectiveness of remedial language teaching methods. Applied Psycholinguistics, 1, 151-170.

[17] Gentner, D. (1982) Why nouns are learned before verbs: linguistic relativity vs. natural partitioning. In S.A. Kuczaj II (Ed.), Language Development. Vol. 2: Language, thought, and culture. Hillsdale, NJ: Erlbaum.

[18] Gleitman, L. (1990) The structural sources of verb meaning. Language Acquisition, 1, 1-55.

[19] Goldin-Meadow, S., Seligman, M.& Gelman, R.(1976) Language in the two-year-old. Cognition, 4, 189-201

[20] Gropen, J., Pinker, S., Hollander, M. & Goldberg, R. (1991) Syntax and semantics in the acquisition locative verbs. Journal of Child Language, 18, 115-151.

[21] Hansson, K. & Nettlebladt, U. (1995) Grammatical characteristics of Swedish children with SLI. Journ of Speech and Hearing Research, 38, 589-598.

[22] Ham, W. (1986) Facilitating acquisition of S-V utterances in children: actions, animations and pictur Journal for Computer Users in Speech and Hearing, 2, 95-101.

[23] Huttenlocher, J., Smiley, P. & Charney, R. (1983) Emergence of action categories in the child: evide from verb meanings. Psychological Review, 90, 72-93.

[24] Hyams, N. (1986) Language Acquisition and the Theory of Parameters. Cambridge, Mass: The MIT Press

Kelly, D. J. (1998) A clinical synthesis of the "late talker" literature: Implications for service delivery. Language, Speech, and Hearing Services in the Schools, 29, 76-84.

Kouri, T.A. (1994) Lexical comprehension in young children with developmental delays. American Journal of Speech-Language Pathology, 3, 79-87.

Kucera, H. & Francis, W.N. (1967) Computational Analysis of Present-Day American English. Providence, RI: Brown University Press.

Leonard, L.B. (1981) Facilitating language skills in children with specific language impairment: a review. Applied Psycholinguistics, 2, 89-118.

Leonard, L.B. (1989) Language learnability and specific language impairment in children. Applied Psycholinguistics, 10, 179-202.

Leonard, L.B. (1995) Functional categories in the grammars of children with specific language impairment. Journal of Speech and Hearing Research, 38, 1270-1283.

Leonard, L.B. (1998) Children With Specific Language Impairment. Cambridge, MA: MIT Press.

Leonard, L.B. & Loeb, D.F. (1988) Government-binding theory and some of its applications: a tutorial. Journal of Speech and Hearing Research, 31, 515-524.

Leonard, L.B., Schwartz, R.G., Morris, B. & Chapman, K. (1981) Factors influencing early lexical acquisition: lexical orientation and phonological composition. Child Development, 52, 882-887.

Levin, B. (1994) English Verb Classes and Alternations: A Preliminary Investigation. Chicago, IL: Univ. Chicago Press.

Lightfoot, D. (1991) How to Set Parameters: Arguments from Language Change. Cambridge, MA: MIT Press.

Loeb, D.F. & Leonard, L.B. (1991) Subject case marking and verb morphology in normally developing and specifically language-impaired children. Journal of Speech and Hearing Research, 34, 340-346.

Markman, E.M. (1989) Categorization and Naming in Children: Problems of Induction. Cambridge, MA: Bradford.

McPheeters, D. & Meyer, J. (1992) The importance of sound, animation and variety in software programs designed to teach disabled preschool children. H.J. Murphy (ed.), Technology and Persons with Disabilities. Northridge, CA: California State University, 347-351.

Naigles, L.G. & Kako, E.T. (1993) First contact in verb acquisition: defining a role for syntax. Child Development, 64, 1665-1687.

National Aeronautics and Space Administration (1991) Programmer's Guide to the General Architecture of Intelligent Computer-Aided Training (ICAT) Systems. Houston, Texas: L. B. Johnson Space Center.

Paul, R. (1993) Patterns of development in late talkers: preschool years. Journal of Childhood Communication Disorders, 15, 7-14.

Paul, R. (1996) Clinical implications of the natural history of slow expressive language development. American Journal of Speech-Language Pathology, 5, 5-21.

Paul, R., Spangle-Looney, S. & Dahm, P.S. (1991) Communication and socialization skills at ages 2 and 3 in "late talking" young children. Journal of Speech and Hearing Research, 34, 858-865.

Pinker, S. (1989) Learnability and Cognition: The Acquisition of Argument Structure. Cambridge MA: Bradford Books.

Radford, A. (1990) Syntactic Theory and the Acquisition of English Syntax. Oxford: Basil Blackwell Ltd.

Rescorla, L. (1989) The language development survey: a screening tool for delayed language in toddlers. Journal of Speech and Hearing Disorders, 54, 587-599.

Rescorla, L., Roberts, J. & Dahlsgaard, K. (1997) Late talkers at 2: outcome at age 3. Journal of Speech, Language and Hearing Research, 40, 556-566.

Rice, M.L. (1993) "Don't talk to him; he's weird." A social consequences account of language and social interactions. In A.P. Kaiser & D.B. Gray (Eds.), Enhancing Children's Communication: Research Foundations for Intervention. Communication and Language Intervention Series, Vol. 2. Baltimore: Brookes, pp. 375-395.

Rice, M.L. (1994) Grammatical categories of children with specific language impairments. In R.V. Watkins & M. Rice (Eds.), Specific Language Impairments in Children. Baltimore, MD: Brookes, pp. 69-89.

[50] Rice, M.L. & Bode, J.V. (1993) GAPS in the verb lexicons of children with specific language impai[illegible] First Language, 13, 113-131.
[51] Rice, M.L., Buhr, J.C. & Nemeth, M.J. (1990) Fast mapping word learning abilities of language del[illegible] preschoolers. Journal of Speech and Hearing Disorders, 55, 33-42.
[52] Rice, M., Oetting, J., Marquis, J., Bode, J. & Pae, S. (1994) Frequency of input effects on [illegible] comprehension of children with specific language impairment. Journal of Speech and Hearing Research, 37, 1[illegible] 122..
[53] Rice, M., Wexler, K. & Cleave, P.L. (1995) Specific language impairment as a period of extended optio[illegible] infinitive. Journal of Speech and Hearing Research, 38, 850-863.
[54] Roeper, T. & Williams, E., (eds.) Parameter Setting. Dordrecht, Holland: D. Reidel, 1987.
[55] Scarborough, H. & Dobrich, W. (1990) Development of children with early language delay. Journ[illegible] Speech and Hearing Research, 33, 70-83.
[56] Schery, T.K. & O'Connor, L.C. (1995) Computers as a context for language intervention. In M. Fey, [illegible] Windsor & S. Warren (Eds.), Language Intervention: Preschool through the Elementary Years. Baltimore: B[illegible] 275-314.
[57] Snow, C., Midkiff-Borunda, S., Small, A. & Proctor, A. (1984) Therapy as social interaction: analyzi[illegible] the contexts for language remediation. Topics in Language Disorders, 4, 72-85.
[58] Staab, C. (1983) Language functions elicited by meaningful activities: a new dimension in langu[illegible] programs. Language, Speech and Hearing Services in Schools, 14, 164-170.
[59] Steiner, S. & Larson, V.L. (1991) Integrating microcomputers into language intervention with child[illegible] Topics in Language Disorders, 11, 18-30.
[60] Tomasello, M. & Kruger, A.C. (1992) Joint attention on actions: acquiring verbs in ostensive and n[illegible] ostensive contexts. Journal of Child Language, 19, 311-334.
[61] Tomblin, J.B., Records, N.L., Buckwalter, P., Zhang, X., Smith, E. & O'Brien, M. (1997) Prevale[illegible] specific language impairment in kindergarten children. Journal of Speech, Language and Hearing Research, [illegible] 1245-1260.
[62] van der Lely, H.K.J. (1994) Canonical linking rules: forward versus reverse linking in normally developin[illegible] and specifically language-impaired children. Cognition, 51, 29-72.
[63] Warren, S.F. & Bambara, L.M. (1989) An experimental analysis of milieu language intervention: teachi[illegible] the action-object form. Journal of Speech and Hearing Disorders, 54, 448-461.
[64] Watkins, R.V. (1994) Grammatical challenges for children with specific language impairments. In R.V. Watkins & M.Rice (Eds.), Specific Language Impairments in Children. Baltimore, MD: Brookes, pp. 53-6[illegible]
[65] Watkins, R.V. & Rice, M.L. (1991) Verb particle and preposition acquisition in language-impaired preschoolers. Journal of Speech and Hearing Research, 34, 1130-1141
[66] Watkins, R.V., Rice, M.L. & Moltz, C.C. (1993) Verb use by language-impaired and normally developi[illegible] children. First Language, 13, 133-143.
[67] Wexler, K. (1994) Triggers and Parameter Setting. Paper presented at the Linguistics, Cognitive Science & Childhood Language Disorders Conference, City University of New York, April, 1994.
[68] Wilson, M.S. (1977) Syntax Remediation: A Generative Grammar Approach to Language Developm[illegible] Cambridge, Mass: Educators Publishing Service, Inc., pp. 59-62.
[69] Wilson, M.S. & Fox, B. (1983) Microcomputers: a clinical aid. In Harris Wintz (ed.), Treating La[illegible] Disorders: For Clinicians by Clinicians. Baltimore: University Park Press, pp. 235-248.
[70] Wilson, M.S. & Fox, B. (1986) Microcomputer language assessment, intervention, and enhancem[illegible] J.L. Northern (ed.), The Personal Computer for Speech, Language, and Hearing Professionals. Boston: Little, Brow[illegible] 101-112.

Information Systems

From Theory to Practice: Checking the Accessibility of Assistive Multi-Media Kiosks

Chris Stary, Alex Totter
B J/ment of Business Information Systems, Communications Engineering, University of Linz, Austria, email office@ce.uni-linz.ac.at

Abstract

Ápaper deals with the evaluation of a particular type of assistive technologies, namely public wkmedia kiosk systems that provide user interfaces for different types of users. Access for all' njiires a key characteristic of the software, namely the adaptation to individual user abilities and «á. The evaluation of these systems is a challenging task due to the setting and the scenario of m that can basically be described through non-linear task accomplishment and short-term mction. Although for evaluation and design principles, such as the flexibility of user interfaces, W been defined, the evaluation of user interfaces for all requires special concepts that do not d; capture the particular scenario of use but also focus on the measurement of adaptation in this Med. In this paper we detail the situation context of evaluation, derive the requirements for mesponding measurements and methodologies, and analyze existing techniques for evaluation, in tar far they meet the derived requirements. We finally introduce a practical methodology that is anntly under empirical investigation.

1. Introduction

Multimedia systems become widely used due to the proceeding establishment of the information iciety and the development of an infrastructure for publicly available interactive technologies. i'Acal examples for public multimedia applications are assistive city information kiosk systems, ach as AVANTI (Adaptive and Adaptable Interactions with Multimedia Telecommunication Ipplications) [4], Massively linked pieces of information allow to navigate in publicly available ntworks, such as the Internet, as the underpinning technology. Since publicly accessible jjmation kiosks have to provide access for all different kinds of users, they also have to provide access for elderly people and disabled users, i.e. they have to be assistive technologies. Univer accessible applications actually require features for the recognition of (dis)abilities and th adaptation towards the acquired user characteristics.

The concept User Interfaces for All [11] has been introduced to meet the increasing demand fo structured development and evaluation of assistive technologies. User Interfaces for All rely *adaptation* as the key principle for universally accessible technologies, namely through provi user interfaces for a variety of end users in a flexible way. Design in the context of User Inter for All is understood to focus on user needs and characteristics that should delive th measurements for checking accessability.

Although adaptation is widely used in the course of user interface development and evaluation part of styleguides, e.g. [7], national and international standards, e.g. [2, 6], or directives, e.g. [the traditional understanding of adaptation does not reflect the situation context of using an evaluating assistive applications. In particular, claiming to support a variety of end user wi different abilities, requires a facilitator to accomplish this task and implement this principle [12]. A a consequence, a clear understanding of adaptation has to be developed, before the evaluation accessibility can be addressed from the methodological perspective.

Consequently, we will wrap up fundamentals in understanding adaptation in section 2 before i detail the situation context for evaluating the accessability of assistive technologies that leads t list of requirements from the methodological perspective in section 3. In section 4 we will the analyze the categories of existing techniques for evaluation in software ergonomics, and check th capability to measure adaptation in the situation context of public kiosk systems. In section 5 w will introduce a novel technique for evaluation that is currently utilized in a field test. Section concludes the paper.

2. Understanding Adaptation

A conceptual and empirical analysis of adaptation has led to the following results [10]:

- Terminologically,
    - Adaptation is closely related to individualization - however, sometimes individualiz is used for generating automatic behavior on user models (in contrast to situation-sp behavior)

- Adaptation is either used as a general term for meeting particular user or task requirements, or to denote automatic behavior of technical systems. In the latter case adaptivity and adaptation are used synonymously.
- When distinguishing automatic behavior from user-controlled accommodation, adaptation is sometimes called customization.
- Customization is also used synonymously to adaptation.

- Adaptation (adaptivity) is based on adaptability, naming a set of features to accommodate user needs.
- Adaptation can be achieved through the following features of technical systems:
  - flexibility towards different behaviors
  - parametrization of components to enable flexibility
  - integratability in the sense of openness to other facilities
  - tailorability to be controlled by the user and to add functionality.

These results have been proven empirically through a semantic content analysis [10]. In addition, the following meta-categories have been identified for evaluation items:

- *individual organization of tasks*: It comprises issues of individualization from an *organization* point of view, such as sequencing subtasks according to individual preferences.
- *individual design of interaction*: It addresses issues of individualization from the *technological* point of view, such as the selection of modalities for interaction according to individual abilities (e.g., speech in/output for visually impaired users).
- *individual provision with information*: It captures issues of individualization from the *cognitive* point of view, such as the individually perceived load of information.

Finally, it has been checked whether adaptability or adaptivity correlate to other principles. In fact, 42 (out of a pool of 570 items concerning the major principles of software ergonomics) items that have not been mutually related initially have been found in semantic correspondence with questions assigned to adaptation. The leading related principle to adaptability and adaptivity is controllability, followed by task conformance, error prevention and robustness.

Overall, the diversity of understandings has been investigated. The conceptual as well as empirical results allow to conclude that adaptation has to be understood as a feature of a system that either enables the adaptation towards end user needs in principle (termed *adaptability*), or automatically adjusts the behavior to internal representations and inference mechanisms (termed *adaptivity*). In addition, adaptability and adaptivity are multi-dimensional principles, since they might be considered from a technical, organizational, social or/and cognitive perspective.

3. Scenarios of Use of Public Kiosk Systems

We now generalize from our experiences in the AVANTI-project [4] the requirements for evaluating public kiosk systems. We first will give typical scenarios of use, and then derive requirements for evaluation techniques, also taking into account the levels of processing adaptation information from the concept User Interfaces for All.

Consider a multimedia kiosk system for city information located at public places, such as railway stations, that is also accessible via the Internet. The system enables everybody to look for information related to the town. Typical scenarios of use are:

Prospective planning of a trip: In the course of interaction, the user first asks for information on monuments, and continues browsing multi-media information on selected ones. The user tries to find out additional information about the

- shortest route to visit the most favorable sites selected
- accessibility of the sites as far as the software can provide this information.

Reactive planning during a trip: The user changes his/her plans after having visited already some of the sites along the route initially planned. S/he wants to visit other sites. In addition, the user might be disappointed about what s/he has seen before. S/he wants to express her/his experiences, and might even expect alternative proposals for further visiting.

Emergency case handling: The user needs immediate medical assistance. S/he tries to find the information about the

- shortest route to get help

essibility of the next site for assistance.

These scenarios indicate neither the user of a public kiosk system does follow a predefined path task accomplishment, nor is his/her behavior predictable, e.g. through applying stereotypical profiles. Hence, we have particular characteristics of the users to be supported, the situation are involved in, and the tasks they are assumed to handle (according to their individual and capabilities).

Characteristics of users:

- *functional roles*: tourist, pilgrim, traveling salesman, inhabitant, etc.; In addition, although these roles might not be mutually related, they might influence the situation and task context, as well as the modalities used for interaction (e.g., the selection of language).
- *abilities* concerning perception, cognition and mobility; These characteristics determine not only the use of devices of interaction and modalities, but also the content to be delivered. An example for the latter fact is the provision of mobility impaired persons with information about the possible access to buildings.
- *demographic data*, in particular age and education. Elderly people and kids have to be considered as user groups with specific needs. The education determines not only the interests, but also the capabilities for interaction at the content level as well as at the dialog level (control), e.g., handling hyperlinks for navigation.

Characteristics of situations: In contrast to traditional settings of human-computer interaction, e.g. VDU-workplace in an enterprise, interfacing a public information kiosk system does not cur in the context of work, and thus, is not constrained the same way, in the sense of mplishing routine tasks embedded in certain organizational conditions and regulations.

- However, the activities performed are intended to achieve certain goals, and thus, represent tasks and subtasks.
- In contrast to usual conditions of work, the situation of end users cannot be considered to be stable, with respect to
    - continuity of interaction and

- completion of tasks.

- *Characteristics of tasks:*
    - *coherence:* Subtasks might not correspond to a linear sequence of activities to adi certain user objectives. For instance, a user might think he/she remember a place's n correctly and tries to find related information typing in the name. Since the system not find the proper information entries, the user has to restart the task accompluj using a search engine.
    - *predictability of task accomplishment:* Since the coherence of activities is not given, behavior for task accomplishment cannot be predicted.
    - *regularity:* Since the tasks cannot be considered to be routine ones (users do not perfo these tasks regularly), the accomplishment cannot be triggered.
    - *differences in granularity of results:* Even for the same task the same user might prod different results, depending on the situation context, e.g. informing about versus boo a hotel both scanning data about accommodation.
    - *task completeness:* It cannot be specified in advance, when a task can be considered to completed. The assignment of the status is done by the end user and depends on his situation context.

As a consequence, evaluating adaptability and adaptativity these data have to be considered to part of the situation context. Additional requirements stem from the architecture of adapt systems [12]: User Interface for All are considered to be designed at a semantic, syntax, and lexi layer, where also the adaptability of an application has to be considered. At the semantic layer information content of an application is the crucial factor. It has to be evaluated whether the specified for interactive manipulation as well as the control in- and outputs are required for accomplishment. The syntax layer concerns the structure of the information provided for control and task accomplishment. Certain rules have to be followed to enable consistent and behavior. These rules establish the 'grammar' of the probably individualized 'language' interaction. At the lexical layer the semantic and syntax information is further refined to co data types that have to be provided for interaction (audio, visual, spatial etc.). The actual ch for communication are specified to enable the physical interaction between the application and intended end users. Evaluation techniques should therefore reflect the structured approach developing universally accessible applications.

4. Existing Techniques

The analysis of existing techniques for evaluation has shown that there exists three different types of evaluation techniques [9]; technology-oriented, user-oriented, and organization-oriented ones. Moreover, traditional evaluation involving end users is based on VDU-workplaces in organizations. Technology-oriented techniques relate measurements to the functionality of an application. ...mental inputs are provided through (inter)national standards and federal directives for the ... of workplaces, such as [5]. In some cases, primarily function-oriented principles, such as ...ency, have been coupled with organizational issues, such as global business processes.

User-oriented techniques emphasize the usability of technological artifacts as users perceive them. Measurements are related directly to human characteristics, such as the blood pressure of users. In contrast to technology-oriented approaches user-oriented techniques do not focus on the functionality of software itself, but on the impact and implications of the functions on the well-being of end users.

Organization-oriented evaluation techniques also focus on human-oriented measurements, such as conditions at work that do or do not support well-being, but not on individual end users. They ... reliable and valid results through checking conditions of work based on standardized descriptions of workplaces, as workflow engines are based on functional roles representing user profiles that are not adapted towards individual end users.

Techniques have been assigned to the dimensions above. It turned out that most of the currently available techniques are either product-oriented, such as EVADIS II [8], completely user-oriented, such as MUSiC [1], or completely directed towards organizational issues, such as KABA [3]. However, in order to evaluate adaptability and adaptivity, all three perspectives have to be integrated, since tasks at public multimedia kiosk systems are exclusively performed depending on user abilities and needs (both may vary significantly), and the context of the situation. The latter influences the individual organization of tasks, whereas the abilities and needs address the user-oriented as well as the technical dimension.

5. Accessability through Adaptation: A Comprehensive Evaluation Approach

AthA (Accessability through Adaptation) is a process-driven approach that involves the end users totally and requires a set of context-sensitive activities to be followed along the different phases of evaluation:

1. for *preparation*; activities to
   - describe the technical system to be evaluated and the measurements for adaptation
   - acquire and specify user requirements with respect to adaptation
   - consolidate the previously acquired knowledge, namely, on one hand, to find the technical features that are candidates to meet the user requirements, and, on the other hand, to identify which measurements can be assigned for evaluation to the relationships
   - select the methodology and methods for evaluation, comprising the identification of specific availability, the tools, and the material to be utilized, and scheduling the evaluation
   - prepare and instantiate the setting for evaluation from the content, organization and social perspective.
2. for *execution*; activities to
   - acquire the raw data
   - monitor, supervise and guide the participants.
3. for *analysis*; activities to
   - compile the raw data
   - analyze the compilation
   - interpret the results of the analysis
   - suggest improvements for end users, developers, evaluators, and coordinator.

AthA meets the listed requirements, as the preparation phase serves for the specification of characteristics of tasks as well as users, including their requirements for adaptation and adaptivity. Hence, on one hand the eomplexity of tasks as well as the constraints for completion can be specified, and, on the other hand, particular user communities, such as elderlies, can be selected as particular target groups for evaluation. From these inputs, requirements for adaptability and adaptivity can be derived. This procedure of customizing the evaluation is continued with (i) task-based specification of scenarios through end users, and (ii) the preparation of methods through the evaluator based on the documented user inputs concerning their expectations.

For instance, the customization procedure for the visit scenario in section 3 might lead to the following items of a questionnaire measuring the three different dimensions of adaptability and adaptivity:

- *organizational perspective*: The navigation through tourist sites allows to choose among several [illegible] structures.
- *cognitive perspective*: The format of the information concerning tourist sites is displayed as the user prefers.
- *technological perspective*: The user is enabled to modify the navigation structure through selection between speech in/output and menu-driven interaction.

In the course of evaluation it is of importance, that the user characteristics are kept constant throughout the scenarios, but the interpretation and perception of tasks to be performed within each of the scenarios are varied. This procedure is required to ensure the quality of results, since the moderator variable are kept constant.

Finally, the requirements given through the application of the concept User Interface for All, are met, since (i) the semantic layer is addressed through the task specification in the situation context of the end user, and (ii) the syntactic layer is addressed through the specification of the capabilities of end users influencing the structure of information to be provided for control and manipulation. Both layers are handled in the preparation phase.

6 Conclusions

Our objectives were to elaborate the specific situation for evaluating public assistive multimedia kiosk systems. We have developed requirements out of the typical characteristics of tasks, users, and the situations where users accomplish their tasks. Based on an analysis of traditional techniques for evaluation that identified shortcoming in meeting the requirements we have introduced ***AthA*** (Accessibility through Adaptation), a situation-aware and process-driven methodology for [illegible], implementing the principle of total user involvement. ***AthA*** is currently under empirical investigation in the context of the AVANTI project, a European initiative to support different user groups when interacting with publicly accessible multimedia kiosks.

7. References

[1] Corbett, M.; Macleod, M.; Kelly, M.: Quantitative Usability Evaluation - The ESPRIT MUSiC Project, in: Human-Computer Interaction: Applications and Case Studies, Proceedings of the Fifth International Conference on Human-Computer Interaction, Orlando, Elsevier, Amsterdam, pp. 313 - 318, 1993

[2] DIN 66234 part 8: VDU-Workplaces, Fundamentals for Dialogue Design (in German), Beuth, Berlin, 1988.

[3] Dunckel, H. ; Volpert, W.; Zölch, M.; Kreutner, U.; Pleiss, C.; Hennes, K.: Contrastive Task Analysis (in German) The KABA-Guide: Fundamentals and Manual, Teubner, Stuttgart, 1993.

[4] Emiliani, P.E.: Access to information in the Information Society. In: Design of Computing Systems: Cognitive Consideration, Proceedings of the Seventh International Conference on Human-Computer Interaction, San Francisco, pp. 481-484, 1997.

[5] EU-Directive 90/270/EEC: Human-Computer Interface. Occupational Health and Safety for VDU-Work (Fifth Directive, Art. 16 Par. 1 of the Directive 89/391/EEC). In: EU-Bulletin , Vol. 33, L 156, Minimal Standard (Art. 4(3) Par. 3, p. 18 , 21.06.1990.

[6] ISO 9241 Part 10: Ergonomic Dialogue Design Criteria, Version 3, Commitee Draft, December, 1990.

[7] Microsoft Cooperation: The Window Interface Guidelines for Software Design. Microsoft Press, 1995.

[8] Oppermann, B. Murchner, H. Reiterer, M. Koch: Ergonomic Evaluation. The Guide EVADIS II (in German), 2nd edition, de Gruyter, Berlin, 1992.

[9] Stary, Ch. T. Riesenecker-Caba, J. Flecker: Implementing the Directive for VDU-Work: The European State of the Art", in: Behavior and Information Technology, in press, 1998.

[10] Stary, Ch. & A. Totter: How to Integrate Concepts for the Design and the Evaluation of Adaptable and Adaptive User Interface, in: Proceedings ERCIM, European Workshop on User Interfaces for All, November 1997.

[11] Stephanidis, C.: Towards User Interfaces for All: Some Critical Issues, in: Proceedings HCI'95, Tokyo, pp. [illegible]-142, Elsevier Science, 1995.

[12] Stephanidis, C.; Akouminiakis, D.: Designing User Interfaces for all Users: Contributions from Applied Ergonomics and Human Factors, in: Proceedings ICAE'96, Istanbul, pp. 137-142, 1996.

DEVELOPMENT EVALUATION OF A WEB-BASED INFORMATION SERVICE FOR AND ABOUT PEOPLE WITH SPECIAL NEEDS

Daniela Busse and Stephen Brewster

Abstract

Need for an information service for and about people with special needs (PWSN) at Glasgow University motivated this case study. The implementation of the system on the World Wide Web (WWW) took place within the framework of iterative and user-centred development, and paid special attention to universal and extra-ordinary design principles. Thorough evaluation, which represents a vital aspect in iterative needs analysis, concerned usability, efficiency of the information design, and meeting of the stated functional requirements. The findings informed subsequent refinements of the information service prototype. In conclusion, the implemented information service was found to meet the high priority requirements stated at the outset of the project. The success of the case study underpins our argument for the importance of the described development and design approach for systems for PWSN.

Introduction

In recent years, there has been a rapid growth in the number of on-line information services, which bring information to the user instead of the user having to travel to the information source [5]. Whilst accessing information resources from a computer at home is convenient for most people, this possibility provides an entirely new opportunity for PWSN. Design of these information sources must take special usability needs as the starting point and not as an added extra. This will also benefit 'mainstream' users [6] [5] [7] [10]. According to Paciello [10], 'assistive and adaptive technology have become the norm while accessible design has become the usability orphan'. In order to achieve accessible design, the Human-Computer Interaction (HCI) aspects of the development process of such information sources assumes a particularly important role. The a this paper is to highlight the development and design framework necessary for 'univ accessible' design. The development evaluation of a Web-based information service for andi students with special needs at Glasgow University will serve as a case study.

2. The Case Study

In the academic session 1996/97 over 700 students were registered as having special ne Glasgow University. However, support facilities at Glasgow University leave a great deal desired. Furthermore, information concerning the range and nature of the few existing facil not being distributed sufficiently. Currently, the Special Needs Adviser (SNA) acts as the 'd of facilities and services. Bits of relevant information are available only in the form of brief distributed by various support agents within the university. These leaflets do not take spe information processing needs of PWSN into account, and several of the support agents are not easily accessible (for instance, the Disabled Students' Interest Convenor (DSIC) Office is sit on the first floor in a building without lift). Thus, a comprehensive, accessible information for and about PWSN was urgently required.

This information service must enable students to benefit from the support available. information must also be designed to make staff and students more *aware* of the charac special needs. Especially for 'hidden' special needs (such as deafness or dyslexia), awareness often becomes the *source* of problems PWSN have to deal with [3]. Thus, we aim

- providing information about existing support for students with special needs;
- raising awareness in the general student/staff body at the university;
- designing the information to maximise communicational efficiency;
- designing the structure of the site to reflect the extra-ordinary HCI needs of the end-users

3. The Design and Development Framework

Iterative development represents a software lifecycle model that emphasises the cyclic, re-i nature of system development [11] [12]. The establishment of the information service descr

This paper pilots the use of this specific methodology in systems for PWSN. The concept of user-centred development formed the framework for this approach. Thus, needs analysis and requirement elicitation (key concepts of the iterative, user-centred methodology) were accomplished in close collaboration with various experts and end-user populations, and pervaded the whole *development evaluation process* of the website.

Extra-Ordinary Design

A significant portion of our population has impairments that reduce the ability to effectively or safely use standard consumer products [13] [7]. Extra-ordinary needs can be categorized into four major groups: visual, hearing, physical, and cognitive/language impairment. Acessible, extra-ordinary design can help mitigating the difficulties experienced by PWSN. Applying extra-ordinary design, the design of the user interface and the structure of the information presented optimally accommodates PWSN's extra-ordinary information processing needs.

Special attention was directed at dyslexia. This is the most common disability among students registered as having special needs at Glasgow University. Dyslexia-friendly design also incorporates the special needs of for instance cognitive impairment. For instance, simple, intuitive content structure, consistency, and avoidance of complex sentence structures will accommodate the needs of both kinds of impairment [4] [13]. The requirements of visual and hearing impairment and of students falling into the broad category of mobility difficulties were also taken into account as far as possible.

Universal Design

Universal design aims for 'usability by all people, to the greatest extent possible, without the need for adaptation or specialized design' [13]. It thus aims at finding a *common denominator* within the requirements of differing user populations.

As Newell [9] pointed out, 'extra-ordinary needs are only exaggerated ordinary needs ... a different point on the continuum of human ability'. He maintained that 'very little HCI research fully takes into account the individual differences among potential users and the vast range of abilities of human beings in general'. Taking special needs into account by applying extra-ordinary design is likely to be also of considerable value to able-bodied people. In the case study, the diversity of categories of special needs and variety of requirements within these necessitated several compromises (see section 4.2).

3.3.User centred Design

Concurrent to the iterative development process, aspects of the system were evaluated from the perspectives of domain experts, and naive as well as special needs end users. In order to assess usability, and functionality, several participatory techniques [11] were employed. These included informal peer evaluation and think-aloud exercises. A field trial evaluation of the prototype was held, including a comprehensive usability questionnaire. Expert walkthroughs included dyslexic students, HCI experts, and the SNA.

4.Universal User Interface Design

4.1.Design Guidelines

Before embarking on the interface design for the information system, several of the published guidelines on general as well as extra-ordinary interface design [6] [1] were referred to. The guidelines typically did not consider a *range* of special needs. Thus, we established some general guidelines by collating those proposed by the consulted sources. The derived high-priority general guidelines were as follows:

- Be consistent within and across pages (for instance location of navigation bar);
- minimize download latency; file size (e.g. avoid graphics);
- assume small screen size and average browser facilities (e.g. no frames);
- provide clear navigational aids (such as clear icons and Table of Contents);
- provide extensive on-line help (accessible from every page);
- succinct page content; avoid burying of information;
- clear labeling of links and pages (e.g. intuitive titles);
- avoid unfamiliar terminology and complicated language.

Reading and writing difficulties are not the only problems dyslexic users have to deal with, as is [illegible] assumed [3]. But also, for instance, comprehension, eye sensitivity, hand-eye coordination, and learning times have to be considered [4]. These principles do in fact promise improved usability for 'able-bodied' users, in line with the universal design paradigm. Therefore, during the cycle of needs analysis and evaluation, the following refined design guidelines were (i)nduced:

- avoid complicated hierarchies of documents by clearly categorizing the sections and by avoiding redundant links between the pages;
- minimize dyslexics' tracking problems (re-reading and skipping lines of text) by using larger text size and a generous amount of white space;
- employ redundancy in information presentation (e.g. hot-text and icons);
- design and layout of buttons must take into account the effects of poor hand-eye co-ordination (e.g. large, clearly separated buttons);
- keep eye sensitivity difficulties to a minimum by avoiding a glaring white background.

III. Design Trade-offs

Dyslexics' eye sensitivity posed a design problem where some common denominator had first to be found. Thus, black text on clear, white background proves most beneficial for partially sighted users, but poses problems to dyslexic ones [8]. The chosen trade-off involved using an 'Old Lace' background, and black text. Hyperlinks, on the other hand, were presented in a moss green colour (see also Figure 1 and 2). Links, especially those comprising the navigation bar, are of crucial importance to successful navigation of the site. Underlined links pose further difficulties for dyslexic users, as the lower part of letters might be concealed and thus lead to reduced readability. To side-step further dilemma for dyslexic users, a black text/white background combination for hyperlinks was thus avoided. We did *not* consider removing underlining of links since screen readers as used by visually impaired people depend on underlining in order to recognize hyper-links.

Another example of a trade-off was the use of icons for navigation. Providing clickable, clear icons with accompanying non-underlined text is the most beneficial design for dyslexics [2]. However, screen readers can 'read' text out loud but cannot read graphics. Screen readers have ere radical improvement for blind and visually impaired people, but unless alternatives to grap made available, they 'may once again experience old barriers they thought had been remove An alternative was to accompany hot-text by icons. However, overabundance of icons (especi they are large and well separated) easily results in cluttered screens. Therefore, navigatio were not to be included. Hand-eye coordination difficulties were taken into account by accompanying hot-text within the page body by easily clickable, large, and well-separated butt

The usability implications of these trade-off design decision were evaluated thoroughly thro the case study. In the final evaluation (section 5.4), mostly positive responses were return particular the nature of the navigation bar and buttons achieved high scores. Navigation thro the site was shown to be easily manageable and to have an intuitive feel.

5.The Development Evaluation Process

5.1.Needs Analysis

Needs analysis covered various perspectives on the system. Techniques employed included interviews, discussions, and walkthroughs, a focus group, and questionnaire surveys. The analysis developed as an ongoing investigation of the problem space concurrent to develo evaluation of the information service.

The experts interviewed in the needs analysis stage were the SNA and the two DSICs. Additi a focus group meeting was organised, which 16 students with special needs attende discussion was vital in highlighting shortcomings in the current promotion of support. problems due to low levels of awareness facing students with special needs were pointed out

An attitude survey was carried out on a representative sample of 74 Glasgow University stu order to assess the levels of awareness in the general student body. In particular, awarenes variety of special needs occurring at Glasgow University, of misconceptions concerning needs, of problems that students with special needs encounter in their university life, and different kinds of support available for students with special needs at university was tested for results indicated a rather low level of awareness in the subjects. This matched the description

...ation as obtained in the focus group discussion and the semi-structured expert interviews (for a detailed discussion of this survey see [3]).

The emerging structure of the prototype thus reflected these different information needs. Three main chapters were created:

- "What kind of support is available for people with special needs?"
- "General information on special needs, and related issues"
- "A Note for tutors, staff, and Advisers of Study".

The first chapter presented the kinds of support available at Glasgow University. It included sections for example on the Special Needs Advisor and Disabled car parks. The second chapter focused on educating the reader about special needs. In particular the problems encountered by PWSN and special needs case studies, including dyslexia, were presented. The third chapter was aimed at staff and contained, for example, notes on staff awareness training.

5. Concurrent Evaluation

Informal peer evaluation took place throughout the development of the information service. Suggestions for improving the prototype often concerned the wording of section headings. The lack of graphics, especially as a background feature was commented on. Graphics were used sparsely in the prototype as they would increase download time, and especially not as background to the pages, as this would have contradicted extra-ordinary design principles in respect to dyslexic and visually impaired users as described above.

Three single person think aloud sessions were conducted. Each of them was designed differently, with respect to the different perspectives taken to the system. Additionally, each of the participants was asked to complete a usability questionnaire. The main concern raised in the first think-aloud session was related to the wording of the main section headings, which proved neither intuitive nor helpful to the participant. The wording of ambiguous section headings was subsequently modified and extended. In general, the system was found 'easy to use', it's structure 'intuitive', and the information 'easy to follow'. She found the educational section interesting and quite clear. The major problem in the second session was the lack of familiarity with hypertext as opposed to linear information structures. The tree-shaped organisation of the data with several access points to choose fa confused the participant. This insight helped to put the design process of the system with resp novice users back into perspective. The main section headings were revised to refer to 'The te main chapters', in order to approach the book metaphor. Additionally, page layout design changed to adhere to two different types: either option or information pages (section 5.3).

The participant of the third session was the SNA. Since she was not a 'naive' subject, i.e. she kn in advance what information the site would contain and how it was structured, the discussion mai focused on content issues. Overall, she was satisfied with the layout, content coverage, and structu of the site. As part of further expert walkthroughs, four dyslexic students commented on the site usability value and completed the usability questionnaire. Mostly positive comments were giv concerning the lack of 'fancy graphics', the font size and colour, the organisation of the site, kind of buttons, the style of language, and the content covered. Another walkthrough involva HCI expert. He rated the section titles self-explanatory, the page content succinct, and fel performance in terms of speed of link traversal highly satisfactory. The function, layout, format position of the navigation bar itself were discussed in detail.

5.3. Major Design Decisions Taken in Response to the Evaluation

Through needs analysis and concurrent evaluation, the initial requirements were refined a expanded. Priorities were laid out for content and user interface design. For instance, the promotia of the SNA, and clarity and consistency in the user interface were seen as crucial.

The requirements listed in Section 4 were taken into account in the design of the user interface the information service. The effectiveness of the measures taken was evaluated throughout. This to further refinements of these requirements. For instance, font size was enlarged. The labelling of links and pages was adapted to the users' needs. Prompted by evaluators' difficulties, italic font avoided throughout the site.

We used a page skeleton to provide structural consistency. This, however, illustrates the importance of the initial design stages within the development process. A modification of the skeleton would force an update on all other pages. Introducing design quality to a system in the later stages mea exponentially more effort having to be undertaken on part of the developer.

inspired by the results of the think aloud protocols, the transition from linear to hyper text by novice users was facilitated by splitting the pages into two groups: option pages (containing only links, i.e. options), which were to be utilized as subroots for further exploration, and information pages, constituting the leaf nodes to this sub-root of the tree hierarchy. An example of both page styles is given in Figure 1 and 2.

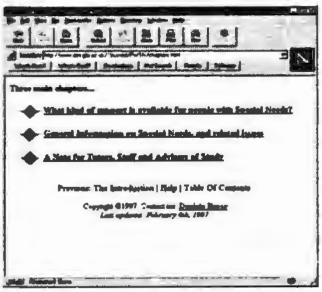


Figure 1 - *Option page.*

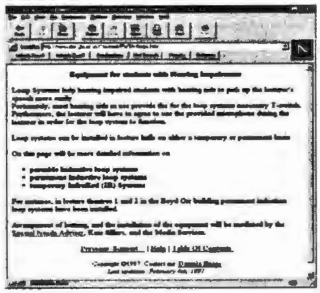


Figure 2 - *Information Page.*

5.4 Final Evaluation

Further design walkthroughs with HCI and Dyslexia experts took place. Additionally, a field trial evaluation was conducted. The trials took place within several departments' computer labs. These were fitted with varying computing facilities, and thus a consensus of opinions across platforms was obtained. Subjects were asked to spend roughly half an hour browsing through the site. Subsequently, they were given a usability questionnaire. Most of the participants found the layout of the pages 'well designed' and 'clear'. Small screens posed a problem, however. Although the site was designed to fit small screens, the scroll bar occasionally had to be used to an extent that was not deemed desirable.

A second, comparative, attitude survey was carried out. It attempted to establish whether subjects' concept of (students with) special needs had changed after having familiarised themselves with the information presented in the prototype. Changes in awareness of the situation of PWSN, and of the characteristics of special needs themselves were found to be significant. In particular, this held true for 'hidden' disabilities. Thus, the data show success in most aspects of the information service investigated.

6.Conclusion

The iterative development evaluation approach taken in the realisation of the information sen has proven to enhance requirements elicitation and prototyping considerably. The nature of t problem did invite such a development perspective. The unclear start and goal states of the syst necessitated that the information that was to be incorporated in the web site be collected an structured in an iterative fashion. Only iterative system development could ensure that requireme were met. The cyclic aspect of the development process also allowed the user interface design pass through several stages of refinement. These ensured high usability scores as obtained in final evaluation studies, and thus met a major aim in the system design. High usability scores gi by the (predominantly) able-bodied users evaluating a system developed in the framework of sp HCI needs guidelines demonstrates the value of universal and extra-ordinary design.

7.References

[1] Accessible Web Design Guidelines; current at 10.1.1998; collection at URL http://www.dcs.jla.« -bussed/MyBookmarks.html.

[2] BOWMAN, L., Dyslexia Archive, email correspondence with McMurtrie, G., 1995

[3] BUSSE, D., Shifts in Level of Awareness of PWSN at Glasgow University, Undergraduate Thesis, Psychology Department, Glasgow University, 1997.

[4] CLARKE, E., Human-Computer Interaction and Dyslexia; at URL http://www.futurenet.co.uk/charity /ado/index.html, 1995.

[5] COOMBS, N., Interfacing Online Services, Alternative Inputs, and Redundant Displays, in [6].

[6] EDWARDS, A. (Ed.), Extra-Ordinary Human-Computing Interaction - Interfaces for Users with Disabilities. Cambridge University Press, 1995.

[7] HOWEY, K. Technology for People with Special Needs: HCI Design Issues, in [6].

[8] IRLEN, H., Reading by the Colour, Dyslexia 2000 Network; site current at 10.1.1998; at URL http://www.futurenet.co.uk/charity/ado/index.html, 1995.

[9] NEWELL, A.F., Extra-Ordinary Human-Computer Interaction, in [6].

[10] PACIELLO, M., Designing the World Wide Web for People with Disabilities: a User Centred Design Appro in Assets '96, ACM SIGCAPH, p. 97, 1996.

[11] PREECE, J., Human Computer Interaction, Prentice Hall, 1995.

[12] SOMMERVILLE, I., Software Engineering, Addison-Wesley Publishers Ltd, 1992.

[13] TRACE, site current at 10.1.1998; at URL http://www.trace.wisc.edu/.

Session VII

Mathematics

LABRADOOR - a contribution to making mathematics accessible for the blind

Mario Batusic, Klaus Miesenberger, Bernhard Stöger

Johannes Kepler University Linz (A), Modellversuch Informatik für Blinde
Altenbergerstr. 69, Austria, phone: +43 732 2468--9232, fax: -9322
e-mail: math@mvblind.uni-linz.ac.at

Abstract

LABRADOOR, a program to present mathematical documents written in LaTeX source code in a Braille maths notation (Marburger Mathematikschrift), is introduced. The program can be used for production of layouted printouts as well as structured electronic documents.

1. Access to Maths-documents

The past two decades saw incredible advances in blind people's access to information technology and in access to written documents. This has substantially improved the situation of the blind in many areas of private and professional life. Access to *mathematics*, however, is still a problem: Mathematical formulae are written in graphical, non linear format which causes problems to the blind. The problem is particularly urgent because blind people, although from their mental possibilities well able to understand and execute mathematical tasks, very often are hindered from success in higher education by the problems outlined: Although mathematics is the most abstract science by making itself 'blind' for the real world by setting up abstract models and acting in these models where the contents never truly rely on the visual sense, its presentation at school and at the university almost always makes use of visual means, thus excluding the blind to a large extent.

Several research activities are going on to overcome these problems, and a lot of progress has been made already especially in the field of audio access to mathematics, e.g.: [1, 3, 5, 6, 12, 13, 14, 15].

Systems for producing a proper layout with mathematical formulae, e.g. LaTeX as well as hypertext systems (SGML, HTML) describe the graphical information in a way which can be interpreted and represented in a structured way for electronic reading and print output. Beside these very promising advances, a lot of problems remain: Mathematical publications are still not really accessible to blind people, nor can the blind presently utilise off-the-shelf computer algebra systems to be aided in manipulating formulae by the computer. Maths documents are often inaccessible because most OCR software which can handle textual material quite well is unable to interpret mathematical formulae. The use of traditional Braille writing equipment to do mathematical tasks is tremendously cumbersome for a blind person, from which an urgent need for an assistive computer software follows. The production of formulae and performing complicated calculations with a computer for the blind is still a very cumbersome task.

The LABRADOOR software introduced in this paper is part of the approach of the working group at the University of Linz to the maths access problem [2, 3, 9, 10, 11]. LABRADOOR is planned and developed as a module which should help us in our researches and tests for a computer supported mathematics working environment. It gives access to mathematical sources by representing LaTeX documents in a format which suits the needs of

- a computer supported working environment
- the computer supported production of Braille outputs and
- the production of electronic documents.

Although LABRADOOR was planned primarily as a module for research on a computer supported working environment, the practical application showed that LABRADOOR as a 'stand alone' program can be used for the production of Braille outputs and electronic documents. Due to the fact that access to documents is the most important part of the support centre at our university and because of our involvement in the production of books for blind students at secondary school level in Austria we put efforts into the use of LABRADOOR for these two tasks.

It should be pointed out that LABRADOOR is not 'an additional solution to the maths problem'. LABRADOOR as it is presented here is only a tool which supports the preparation of maths documents for the three tasks mentioned, and therefore a preparation tool for possible 'solutions'. LABRADOOR should also not be seen as a new maths notation: It should act as a bridge between computer oriented notations like the widely used LaTeX, and blind user oriented codes such as

Marburg notation. We think that building bridges between computer oriented notations and Braille notations are a prerequisite for further developments of a user friendly access to printed or electronic documents and a computer supported working environment for mathematics.[1]

1 LABRADOOR - General description

LABRADOOR acts as a LaTeX-to-Braille door. LaTeX is a system to produce and to layout texts with mathematical formulae in excellent typesetting quality [8]. The program got extremely widely used for the production of mathematical and scientific publications. Therefore, a door to LaTeX documents can be seen as a common basis for mathematics especially at universities. Although LaTeX can be read directly by the blind it does not exactly fit their needs because of disadvantages in reading LaTeX (and also other notations like MathML) sources in Braille [2, 3]. A first step used was the 'Human Readable TeX' (HRT) which offers better access to LaTeX sources for direct use. As mentioned before, the goal of the research should not be seen in such a new, additional notation. Although HRT is readable for the blind there are notations available which have been optimised for Braille reading over decades and which better fit the needs of the readers of mathematical formulae. In the German speaking area, the *Internationale Mathematikschrift für Blinde* [6], also called "*Marburg Notation*", is a notation which, although by far not universally used, has gained some popularity among educational institutions for the blind. It can be seen as a suitable and accepted standard for reading mathematics in this area. Acceptance of documents and products for the blind in the field of mathematics will be enhanced if this standard is supported.

Some remarks about the advantages of the Marburg Notation for the blind reader are in order: First, it is a notation optimised for space, presenting mathematical expressions in a rather compressed format: Since Braille occupies extremely much space both on paper and on a refreshable Braille display, this advantage is of special importance. Second, it is especially close to literal Braille: A student who has learned traditional Braille at school can quite easily advance to the Marburg maths

[1]) The need for such a bridge between computer oriented notations and the standard Braille codes is also documented by the fact that similar projects are planned or run in other countries like France and USA.

notation. On the other hand, the Marburg Notation has some disadvantages that are to be mentio
First, it is a rather old notation which is very seldom updated - the latest revision dates from 19
the latest but one edition from 1955. This implies that it does not reflect recent changes in the
notation of the sighted. Second, its compressed presentation comes at the price of a
complicated syntax - we believe that the syntax of the Marburg Notation needs stri
simplification to keep the system usable.

Nevertheless, there is no (urgent) need for a new, additional notation for sequential readi
mathematics. Since, as mentioned, many mathematical publications are presently available
LaTeX format, it is more urgent *to open a door* towards their conversion into a suitable and
known format. LABRADOOR should be seen as a platform for access to LaTeX docume
notations which are known and used by the blind. As a first solution for the German speaking
LABRADOOR translates LaTeX into grade one Braille and thereby the mathematical part
LaTeX into Marburg Notation. LABRADOOR as an interface to LaTeX fulfils the needs
different tasks in our practical work of producing printed and/or electronic books and material
the blind.[2] Especially since one year, LABRADOOR is used in practice to produce mathe
literature in Braille for students at secondary as well as university level in Austria. LaTeX als
the needs of producing well layouted ink print versions from the same sources. LABRADO
the abstract layout description of LaTeX for producing Braille-layout for

- a printed output or
- an electronic version

in Braille. In the following we give an overview of the functions of LABRADOOR as a layout
for Braille and electronic documents.

---

[2]) To support the production and the reorganization of LaTeX sources, a special LaTeX Editor, developed by
working group, is used. Most documents contain a lot of macros which have to be resolved. This editor is also u
the sighted, especially by teachers who teach blind students for producing documents for the blind as well as for
print output. LaTeX and Braille beginners can start with their job without doing a lot of studies before. Since the
contains templates for structuring a document and for typesetting mathematical expressions, common mistakes
avoided. These documents can be handled by LABRADOOR which can be called out of the editor. You can call
TeX for producing a black print output or a screen preview of the material to proof ist correctness. In the followi
shall concentrate exclusively on LABRADOOR and its functions for producing Braille print or electronic docum
Braille. We start from already available LaTeX documents without any user defined macros.

LABRADOOR - Basic Features

3.1 Input

The documents treated by LABRADOOR have to fit the TeX–LaTeX conventions concerning the organisation of input data and mark–up language. Besides the elementary input conventions described in the LaTeX Book, LABRADOOR supports some packages changing and/or extending input conventions:

- Package *german* — identical to the *german* option of the package *babel* for multilingual document preparation
- *greek* option of the package *babel*
- nearly all options of the package *inputenc*, which allows one to use different code pages (encodings of the upper ASCII area) such as *cp437, cp850, latin1, applemac* etc. The complete listing of implemented input encodings you can find in the file *lbdl.cf* of the current LABRADOOR version.

Each input encoding table has its corresponding derived output table allowing us to produce computer Braille versions of a document for almost all platforms.

3.2 Working Modes

Based upon two fundamental ways of treating input data, we can distinguish between two modes of work within the LABRADOOR program. The greatest part of a document builds mostly the normal text — in words of computer science: sequences of literals, accented literals, punctuation marks and numerals. During all the time in which the input consists of 'pure text', LABRADOOR is working in so–called *text mode*. There are many ways in LaTeX to switch to a mathematical mode. Once within this mode, the input characters have another meaning — they are elements of a formula. On entering such text parts, LABRADOOR changes also to one of several mathematical modes, preparing the material of the formula for the representation in an implemented Braille mathematics notation.

3.3 Text mode

The 'text parts' of a document are handled by the procedure *Text*. All literals, accented lia (being special LaTeX commands), punctuation marks and numerals are mapped accordinj toi output table of the active language. It is also possible for the user to implement new output supporting other languages. The procedure *Text* has also a routine *hyphenate* which carries automatic word hyphenation on line breaks according to the hyphenation patterns of LaTeX for active language. It is presently possible to print according to the output tables derived from implemented input table, such that computer Braille representations on many different platforms supported(about input see section:). Moreover, one can choose German shortwriting grad (*Deutsche Vollschrift*). The user extendible implementation of shorthand systems is planned.^

3.4 Mathematical Mode

If LABRADOOR encounters a formula within a document, it enters the corres mathematical mode. There are currently three implemented maths modes:

- in-line mathematics mode, handling the LaTeX environment beginning with \(and e with \). The formula is placed in the text flow of the current paragraph
- displayed mathematics mode, handling the LaTeX environment beginning with \[ending with \]. The formula is placed in its own line of text
- equation mode, handling the LaTeX environment *equation*. Just like displayed mode, with an additional automatic numbering of equations.

The differences between these modes refer only to the treatment of the formula by LABRADOOR formatter. The main job of the *Math* procedure is to convert the textual encoding of the formula its binary tree representation and subsequently to call a routine which prints the formula acc to the rules of the chosen Braille mathematics notation. Presently, the only implemented mathematics notation is the Marburg Notation.

4. LABRADOOR Layout features

LABRADOOR can interpret the layout description of LaTeX for the production of layouted Braille material. LABRADOOR as a stand alone tool offers a command line oriented interface under DOS. LABRADOOR can be executed with the command *lbd* followed by the name of the TeX file which ends with **.tex*.

Lbd test

LABRADOOR produces an output file which is named according to the used configuration (**.ptr* for a paper output, (**.txt* for an electronic document). The name can be chosen independently by defining him as a command line option

Lbd test result.txt

LABRADOOR uses a configuration file **.cf* to translate LaTeX into Braille. The layout for electronic as well as Braille-print documents can be defined in such configuration files which can be chosen by adding the name of the configuration file at the command line call.

Lbd test result.txt book.cf

In the following you find a selected list of features which are already implemented and which can be defined in the configuration file:

- output medium (computer/printer)
- title page
- table of contents
- page layout (lines per page, characters per line, left/right/bottom/top margin, page number, horizontal/vertical text shift)
- indentation (first line of paragraph, quotation, itemise, enumerate, table of contents)
- tables (automatic column break, space between elements, space handling, horizontal line representation characters)
- headings (characters for underlining level 1-4, table headings, heading for the table of contents; header centring on/off)
- automatic numbering (page number, section levels 1-4, table numbering, footnotes, enumerated lists)
- counter formats (arabic, roman-small, roman-capitals, alphabetical for page number, original page number, section levels 1-4, table numbering, footnotes, enumerated lists)

- itemise characters
- math environment separators (Braille characters to mark the transition from text to ma
mode and vice versa)
- printer commands (ASCII sequences to be sent on the beginning, the end of file, after a
page, ¼)

5. Further development plans

Features like printing of indices, bibliography, and citations will be implemented in the n
months. Other features of LABRADOOR will be implemented in the future:

RTF input: People preparing materials which should also be used by blind persons often pre
direct visual formatting in the document preparation process. Most widely used are various versi
of the MS-Word text processor, mostly running under Windows or on the Apple Macintosh, ma
rarely under DOS. To respond to the wishes of these people, one of several RTF-to-LaTeX
converters should be integrated into LABRADOOR to make RTF also a valid input format in
LABRADOOR. However, we have to emphasise that mathematical formulae even then have to
inserted in LaTeX.

HTML output: On the output side, the possibility of producing HTML documents should b
implemented. A symbiosis between LABRADOOR and existing converters from LaTeX to HTML
for example one by Nikos Drakos [4] should be integrated.

Some further features:

- Integration of semi-graphic (realised by text characters) and real printer graphic
- implementation of some additional Braille mathematics notations besides the Marburg
Notation
- implementation of a universal shorthand module allowing the users to integrate diffe
shorthand systems for multilingual use

References

[1] BARGI RANGIN, H., SAHYUN, S., LUNDQUIST, R., GARDNER, J., TRIANGLE - Tri-Modal Access To Text, léte, Bjuations, Graphs, Calculations, And Figures, in: Proceedings of the 1997 CSUN Conference on Technology Persons with Disabilities, Los Angeles, CA, March 1997

0 BURGER, F., KNASMÜLLER G., MIESENBERGER K., STÖGER, B., Access to mathematics for the blind, in: Bager, D. (Ed.): New technologies in the education of the visually handicapped; Colloques INSERM/John Libbey Eniext LTD, Mountrouge, Paris 1996

P BURGER, F., BATUSIC, M., MIESENBERGER K., STÖGER, B.: Access to mathematics for the blind - Defining HEX Standard, in: Klaus, J., Auf, E., Kremser, W., Zagler, W., L. (Eds): Interdisciplinary Aspects on Computers Helpin People with Special Needs; 5th Interantional Conference, ICCHP '96, Oldenbourg, Wien, München 1996

[DRAKOS, N., The LaTeX 2 HTML Translator, Computer Based Learning Unit, University of Leeds}, URL: ly ljnvw-dsed.lnl.gov/files/programs/unix/latex2html/manual/

[5] ENGELEN, J. J., BAUWENS, B., BORMANS, G. and EVENEPOEL, F., Mathematics for print disabled persons using structured documents, in: Proceedings of the Tenth CSUN Conference on Technology and Disability, California State University at Northridge 1996

f EPHESER, H., POGRANICZNA, D., BRITZ, K., Internationale Mathematikschrift flir Blinde, Deutsche stdienanstalt, Marburg (Lahn) 1992

NER, A., BARGI RANGIN, H., LUNDQUIST, R., BARRY, B., PREDDY, M., SAHYUN, S., SALINAS, , New Methods of Reading, Writing, and Manipulating Information by People with Print Impairments, in: edings of the Fifth International Conference on Computers Helping People with special needs, Linz, Austria, ly, 1996

[8] LAMPORT, L., LaTeX - A Document Preperation System, Addisson-Wesley, Reading, Mas., 1983

PILICHTENBERGER, F., MIESENBERGER, K., STÖGER, B., KNASMÜLLER, G., 1. Forum ‚Zugang Mathematik It Blinde", Universität Linz 1995

[10] LICHTENBERGER, F., MIESENBERGER, K., STÖGER, B., BATUSIC, M., 2. Forum ‚Zugang Mathematik flir Blinde", Universität Linz 1996

[11] LICHTENBERGER, F., MIESENBERGER, K., STÖGER, B., BATUSIC, M., 3. Forum ‚Zugang Mathematik flir Blinde", Universität Linz 1997

[12] RAMAN, T. B.: Audio System for Technical Readings (ASTER), URL: http://www.research.digital.com/CRL/personal/raman/raman.html

[13] STEVENS, R. D., EDWARDS, A. D. N., HARLING, P. A., Access to mathematics for visually disabled students through multi-modal interaction, in: Human-Computer Interaction, Special issue on Multimodal Interfaces, 12 (1&2) pp. 47-92, 1997

[14] WEBER, G., Stevens, R. D.: Integration of Speech and Braille in the MATHS Workstation, in: Klaus, J., Auf, E., Kremser, W., Zagler, W., L. (Eds): Interdisciplinary Aspects on Computers Helpgin People with Special Needs; 5th nternational Conference, ICCHP '96, Oldenbourg, Wien, München 1996

[15] WEBER, G.: Konzepte zum Mathematikunterricht mit Techniken des World-Wide Web; in: Dresdner Kolloquium 1996 Mit der Braillezeile auf die Datenautobahn, http://elvis.inf.tu-dresden.de/koll96/artikel8.html, Dresden 1996

TechRead: A SYSTEM FOR DERIVING BRAILLE AND SPOKEN OUTPUT FROM LaTeX DOCUMENTS

Donal Fitzpatrick & A.I.C. Monaghan

Abstract

One of the most difficult aspects of research for a blind student is the unavailability of technical material in a format accessible to them. To date, much of the effort of transforming documents into either Braille or spoken output has been in the literary rather than in the technical or scientific areas. For example, the majority of the spoken text produced by existing screen access technology does not harness the capabilities of synthetic speech devices but instead outputs the material using a monotone. TechRead, on the other hand, is a system which, it is hoped, will render technical documents more accessible to blind people. This software will take LaTex documents as input and produce Braille or spoken output from them.

The main aims of TechRead are as follows:

- *To take a subset of LaTeX and produce both Braille and spoken output from it.*
- *To accurately model a document and to present this to the blind user using a simple and intuitive interface.*
- *To harness the capabilities of synthetic speech devices to give more meaningful spoken output to the user.*

This paper discusses the fundamental principles underlying this system. It aims to show how the LaTeX document is transformed into an internal representation of the document, and from this to either Braille or spoken output. The final section discusses how the system will be expanded to cater for mathematical material, and our beliefs that the ideas contained in the system can be used to improve screen access technology.

Keywords: LaTeX, Spoken Documents, Braille

1 Introduction

For many years, the focus of those writing software to translate material into Braille has been [illegible] technical documents. Much of the effort has gone into producing software which will translate [illegible] material into Braille while ignoring the more technical documents. Therefore, as can be imagined, the procurement of technical data for those who cannot read the printed version is both a time-consuming and tedious affair.

TechRead aims to solve this problem. The purpose of the system is to take a file in LaTeX [2,3] and produce an output file in some medium accessible to blind people. A two-fold approach is taken. Firstly, the system will take the input source file and derive a Braille file from it. The user will then [illegible] original hardcopy of this file using a standard Braille embosser. The reason for using Braille is simple, many people think that it is an out-dated, archaic system, which has no use in the modern age of advanced technology. However, it is a fact that for many people Braille has been their standard means of reading for all of their lives, and in our view they must be catered for by TechRead. The translator will, as closely as possible, represent the structure of the document i.e., the Brailled document will be as close a replica of the printed version as it is possible or sensible to be. There will, by the very nature of Braille, be some discrepancy here particularly in the area of lay-out.

While Braille to many is the only means of reading a document, it is limited. By its very nature as a tactile form, Braille cannot convey many aspects of documents which sighted people find so important, and which they take for granted. For example, it is only possible to show emphasis in Braille in one way. In order to do this, the emphasised text is italicised no matter what the printed font might be. Speech on the other hand allows for a far wider scope. It is possible for example to have different voice characteristics for emboldened or italicised text. Another major advantage of speech is its ability to convey both content structure and layout characteristics of a document by the use of prosodic characteristics.

TechRead's second mode therefore combines existing speech synthesis technology with an analysis of document markup information to produce a "speaking document browser". Our current strategy is to use the LaTeX as an input source and to produce an off-screen model of the document from this. Using this model the blind person will be able to read a document in as similar a manner to their sighted colleagues as possible. An example of how this system will work is as follows.

Let us assume that the document being browsed is a newspaper, with sections, headlines and articles. The sighted person will read the section if it interests them. However, there might be headlines in this section that they wish to skip, or paragraphs in the articles which they do not wish to read. The document browser aims to provide the blind user with exactly the same functionality. The browser will allow the blind user to skip sections, paragraphs etc. Also, to return to the analogy of the newspaper, if the sighted person wishes to find the next headline, then they can simply scan down the page to see it. The document browser will allow the blind user to go directly to the next/previous section or sub-section of the document.

The advantages of such a system would be many. Unlike the current situation the blind reader would not have to read superfluous and extraneous information. They could "scan" the document using the browser, until the relevant material has been reached and then read it.

One of the key underlying ideas of the TechRead system is that of conveying the structure of a document to the blind user. In order to achieve this goal, an off-screen model (OSM) of the document must be constructed. The strategy employed in producing the accessible documents is based on a three level architecture as shown below (Fig. 1.1). As can be seen, it consists of an input or source file (LaTeX) which is passed into a pre-processor. This pre-processor will then convert this raw LaTeX material into an internal representation of the document, which can then be passed on to either the translator or the [illegible] for producing the documents used by the Browser.

Before embarking on this discussion however, it would be useful to outline the reasons for selecting this type of architecture. Firstly, such a system lends itself to a very modular design. The layered structure means that one component can be changed without altering the overall structure or logic of the entire system. Secondly, though at the time of writing the input source file will be in LaTeX, there is no reason why other file formats cannot be added at a later date. All that will be necessary will be to write the conversion routines to transform the input file into the internal format used by the translator, and generator for browsable documents.

2. Representing structure.

The off-screen model of the document is constructed at the input stage. At this point the structure of document is encoded. For example, to return to the analogy of the newspaper, it is at the input stage translation that the structural information such as the whereabouts of the starting points of sections headlines would be deduced. This system will enable the blind user to browse a document in as manner to a sighted colleague as possible. The off-screen model will enable the reader to do this. The following paragraphs will describe the model used by TechRead, and the manner in which this influence the design of an interface to the document.

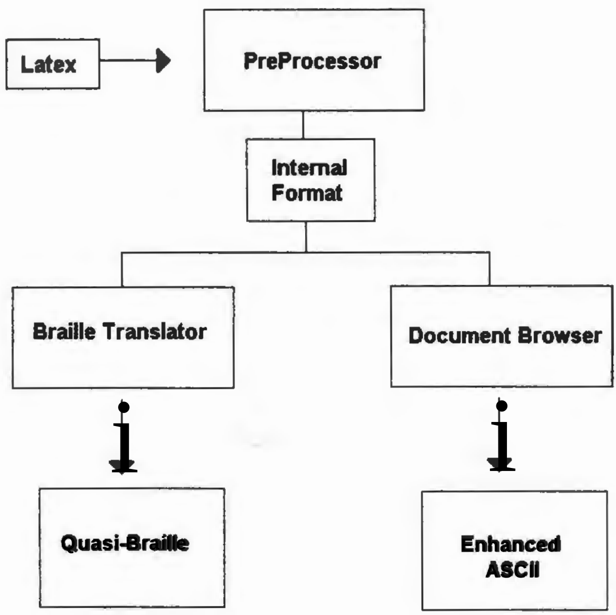


Figure 2.1: The Three Layer Architecture.

Previous systems [7] used a tree based architecture to represent the structure of a document. TechRead uses a complex hierarchical structure to represent this. We assume two types of node in the system, being a terminal while the other is an internal node. The terminal nodes will be used to hold the actual of the document, coupled with any formatting information associated with that text, while the internal nodes will be used to hold the material relating to headings, sub-headings etc. The architecture can best described as a cross-linked tree. The root is a node containing all global formatting for the document. Below this are the first level headings (if they are present), or simple terminal nodes containing the text of the document otherwise. At all levels below the root, the nodes are inter-linked both downwards and across the same level. For example, each section node is linked to the preceding and following sections, as well as dominating the sub-sections contained within it: this allows the user to browse any chosen the document. In addition, the left-most terminal node on any branch of the tree is linked to the left-most terminals on the preceding and following branches: this allows the user to skip forward or back a paragraph. All terminal nodes on any given branch are linked to each other in the form of a list. Finally, the right-most or final terminal on any given branch is linked to the left-most terminal on the next branch for smooth continuous reading. This combination of links directly models a range of different reading strategies.

During construction of the OSM, any formatting changes are passed up the hierarchy to enable rapid processing of the document. For example, if a portion of emphasised text appeared in paragraph 4 of Section 3, a flag would be set in both the section and paragraph nodes. Thereafter, if the browser encounters that section it will examine the paragraph nodes to find the one which contains a formatting command. Similarly, browsing the paragraph level would lead to an examination of the terminal nodes to discover where the formatting change occurs. The algorithm which ultimately produces the spoken version of the document would then alter the characteristics of the voice appropriately, instead of simply outputting the text in the normal reading voice. As a consequence of this model, the interface to the document can be very flexible.

It was initially decided to display the LaTeX on screen in an uninterpreted form and to design the interface to the spoken version of the document such that it was based around the numeric keypad on a standard IBM-compatible computer. This is in keeping with trends in the design of modern screen access technology, where developers attempt to ensure that the time taken by users to learn the system is kept to a minimum. However, use of the numeric keyboard also has an inherent logical basis. Firstly, navigation through the document is intuitively related to the directional keys (up = 8, down = 2, etc.). Secondly, the use of meta-keys in combination with the numeric keypad allows functions at different levels of the document. Let us assume, for example, that the “5” key on this keypad reads the current character. When pressed in conjunction with the “shift” key it could read the current word, and in conjunction with the “control” key could read the current paragraph.. The flexibility of such an interface leaves a wide scope for expansion or customisation. The number of overlays which can be placed on the numeric keypad is (theoretically) infinite, while the fact that only a small number of keys are at the core of the system means that should the user desire to alter the key mappings it will be relatively straightforward to do so.

3 Translation Algorithms

We have seen how the TechRead system takes the raw LaTeX documents and produces an off-screen model from them. Then next phase of producing accessible documents is to transform this model into either speech or Braille. The following sections detail how this will be achieved.

3.1 Braille Translation.

For many years, the means for producing Braille material from various types of document have been known. However, as was stated in section 1, the material produced to date has been of a literary rather than a technical nature. There are still very few translation packages which can take technical documents with embedded mathematical formulae and render accurate Braille. This is one of TechRead's main aims.

The translation process simply involves a character substitution algorithm, consisting of a rule based engine which replaces patterns of characters found in the input document with their grade II Braille equivalents. The rules are of the form:

"input_string" → "Braille symbol"
"input_string" → "Braille_symbol"+ "remaining_input"

Either the entire string has a Braille equivalent, or the first part has a Braille equivalent.

However, more important than the translation process is the actual material which is translated. How will example should emphasis be conveyed to the Braille reader. Traditionally Braille has used only one form of emphasis, namely italics, and this has been used to denote emphasis irrespective of the form of enhancement of the document. This notion of "what" rather than "how" to translate is partic important when dealing with mathematical material. Unlike the spoken version of the document where almost infinitely varied set of alterations in the characteristics of a voice can convey much of the sen of the formulae, Braille has only one way to translate mathematical material[1] One of the means in wh it is hoped that this translator will improve on others currently in existence is the means in which it t spatial location of formulae on a page. For example, as any student using the British math notation will know it is customary to simply write equations in a line across the page (as though it v literary text) instead of using the conventions adopted by typesetters for displaying printed mathem While it is true that not all of these conventions will have relevance to Braille mathematics some of th such as the use of vertical as well as horizontal orientation to display formulae will improve readabilitji

3.2 Producing Spoken Documents

By far the more interesting portion of the system is that concerned with the production of spo documents. While the derivation of Braille from the LaTeX input may indeed be very useful there is more variety in the output which can be obtained by actually using alterations in the characteristics of voice to convey the material to the user.

To begin with, let us consider the text-to-speech synthesisers[2] which are currently in common usage. of them take ASCII characters as input. These characters consist of the text to be spoken coupled some control sequences which can alter the characteristics (such as pitch, voice quality and speaki of the voice. However, some synthesisers are more flexible than others, allowing the programmer gre control of the adjustments which can be made. Accordingly it has been decided to write the docu browser in such a way as to be independent of synthesisers. In order to do this, we will calculate pitch duration at a symbolic level, e.g. duration as very long, long, normal, short, very short. We will determine the correspondence between this symbolic level and the control codes for a partic synthesiser. After the prosodic model has been computed, the appropriate substitutions will be made ensure that the appropriate control codes are sent to the synthesiser at the appropriate time.

Current text-to-speech systems process basic ASCII files and extract information about the structure relative salience of units of text. On the basis of this minimal information, they control the pros aspects of the synthetic voice. It seems obvious that formatted texts and texts with various lev markup will provide much more information on structural and salience relations. These relations t generally expressed in natural speech by a combination of pitch, timing, intensity and voice quality. developing a "speaking document browser", our main task will be to relate markup and format commands to the various possible prosodic realisations.

---

[1] There are several versions of Braille Mathematical notation but within each set there is only one way to translate given formula.

[2] A device which produces spoken output from ASCII text.

[intend to use a prosodic model based on previous work in prosody for speech synthesis [4-6]. lody in TechRead will be represented by abstract symbols which stand for different levels of emphasis, hus hierarchic relations, and paralinguistic effects such as changes in speaker, attitude, and emotion. issume that the large number of formatting commands which are available in LaTeX and similar ument preparation systems actually correspond to a much smaller number of linguistic categories ich can be realised prosodically. These categories include subordination, aside, change of topic, list, phasis. For example, an aside may be encoded in a document as a footnote, a parenthesised ge, a margin note or simply some text between commas; all of these might receive the same odic treatment in a spoken rendition. Similarly, in the spoken version it may not be desirable to guish between bold, italic, underlined, capitalised and quoted text: it seems unrealistic to expect the er to keep all these emphasis types distinct. Moreover, TechRead is limited by the possibilities of the esis devices which will produce the spoken version: not all synthesisers offer the same degree of trol over the prosodic realisation, and the granularity of control also varies. However, previous work has shown that a small number of prosodic categories allows the construction of quite complex thies which should be sufficient to express all the relations which sighted users extract from med documents.

rting with a core set of LaTeX commands, we will derive a model of the possible functions performed ferent formatting commands. Each of these functions will relate a set of formatting commands to a que combination of prosodic symbols. These symbols will be given a translation in terms of the control Bues for each output device. The acoustic phonetic details of the spoken output will therefore pend on the particular synthesiser in use. The core set of LaTeX commands will then be expanded to ude almost all the standard commands [3], although this will always be a subset of full LaTeX: we not cope with user defined macros.

It is hoped to construct a formal model for the alterations in the voice characteristics to show proper mantic interpretation of mathematical equations in particular. It should be noted that to date much of is work relating to this portion of the system has gone in the direction of enhancing the spoken text, as posed to mathematical equations. (For a discussion of our future work in the area of mathematics see An #4.)

[order to translate efficiently from the format used in the OSM to a synthesiser-independent format for poken output, we have devised an algorithm based on the off-screen model generated for each specific %ument. As was said in section 2, flags are stored as part of the internal nodes which indicate whether anges of formatting occur at a lower level The algorithm simply checks the formatting of each level as goes, and, if no change occurs at that level, the text at the levels below this is output to the browser with ditional control codes. However, if a formatting change is detected, the translator drops down a level and goes through the same process, until a point is reached when the text contained in the terminal nodes is found. At this juncture the algorithm simply scrutinises the formatting of the text and, where essary computes the prosodic changes necessary to convey the visual appearance of the material to the lind user. An example will suffice to explain this algorithm.

"us assume that a default reading voice (VI) has been chosen on a DECtalk [1] synthesiser, and that are translating a document of two sections with no sub-sections. The browser encounters the starting int for "section 1" in the OSM, and checks the Mformation stored to determine whether there is a need to alter the voice characteristics within that section. Let us assume that in section one there is no such ration needed. The material contained in this section can now be simply output for use by the document browser. However if in "section 2" there is some emphasised text in the first paragr translator will not now simply output the text, but will discover that a change occurs in the section, so examine the paragraph level. Here it will become apparent that there is a change in the first parag the translator will now examine each of the words within this paragraph to deduce where the alterati voice occurs. When the start of the emphasised text is found, V2 replaces V1 as the reading voice, u the attributes change again, when the voice is returned to V1 for normal reading.

4. Discussion

Though it is intended to incorporate mathematical translation into the TechRead system the majority the work done thus far has been in the realms of conveying the structure and visual enhancem document to the listener. Therefore, as can be seen from previous sections, the algorithms devised to have been for the production of both Brailled and spoken text. However, these algorithms have designed with mathematics in mind, and it is our belief that minor modifications will ensure that this of material will be translated successfully.

It is our intention to conduct a study over the coming months to determine what information sig people extract when reading equations or other formula-based material. We intend to show them a s of mathematical expressions for various fixed lengths of time, and get them to write down what recall. The use of varied lengths of time is intended to simulate different types of reader. For exam is hoped that the replies we get after the subjects have seen the equations for the shortest time will indi what a sighted person sees when they glance at an equation, while those we obtain after the longest of observation will indicate what they recall after examining the mathematical material in depth. It is hoped to analyse these results to determine the best and most effective way to give a listener a "glance" an equation.

Much work has been done in this area to date. The method used in The Maths Project [8], for insta was to convey the "glance" using musical tones. It is hoped to find a more natural alternative to this. currently envisage several reading modes for equations: verbose, overview and glance. In "glance" mo the system will announce the presence of each equation, followed by an indication of the component the equation e.g., "Equation: summation followed by integral followed by fraction". This inform should be available from the formatting commands. TechRead is intended to process equation of the t and complexity encountered in pre-university examinations.

We also foresee uses for this system in the realms of screen access technology. The traditional appro to designing screen reading software has been to start from the operating system, and then design a which will cope with various types of package. It is our belief that TechRead could be adapted to with many different types of documents. For example, a spreadsheet is simply a table, and a docum produced by a word processor is simply a document marked up in different ways. Accordingly, believe that instead of designing screenreaders to cope with various operating systems, it may be poss to incorporate "style sheets" into the TechRead system, thus rendering many different types of docu accessible. Finally, though LaTeX is being used at present, there is no reason why the input source co not be amended to SGML at a later date.

ices:

Talk, a trade mark of Digital Corporation, http://www.ultranet.com/~rongemma/

UTH, Donald E., The TeXbook, Addison-Wesley, 1993

PORT, Leslie, LaTeX: A Document Preparation System, Addison-Wesley, 1986

MONAGHAN, A.I.C., Intonation in a Text-to-Speech Conversion System, Edinburgh University thesis, 1991

ONAGHAN, A.I.C., Intonation Accent Placement in a Concept-to-Dialogue System., dings of AAAI/ESCA/IEEE Conference on Speech Synthesis, New York, mber 1994, pp. 171-174

MONAGHAN, A.I.C. & D.R. LADD, Manipulating Synthetic Intonation for Speaker erisation., ICASSP 1991, vol. 1 pp. 453-456

MAN, T.B., ASTOR: Audio System for Technical Readings, Cornell University Ph.D. thesis, 1994

The Maths Project, http://porta.cs.york.ac.uk/maths/

8-Dot-Braille for Writing, Reading, and Printing Texts which Include Mathematical Characters

Waltraud Schweikhardt[1]

Abstract

In this paper we introduce the Stuttgart mathematical notation for the blind and explain, how blind children, students as well as mathematicians can use it on a computer. The typed characters are immediately represented on the connected braille-display as well as on the screen in a representation, which is understandable for sighted persons. The typed material can be printed for the sighted, and it can be embossed for the blind. We explain the components and the technical requirements of a suited working environment. Finally we point out how existing mathematical materials can or could be translated into a tactile representation.

1. Introduction

Until today, there exist different versions of mathematical notations for the blind, though mathematics for the sighted has an international character. Most of the notations, which are used today, are based on six dots. These are for instance the „Internationale Mathematikschrift für Blinde" [1], the Marburg mathematics notation, which is also known as the Marburg system, and the Nemeth code [2]. In several schools in Germany the „Stuttgarter Mathematikschrift für Blinde" (SMSB) [6, 9] is used. Other notations are under development [3]. Another approach to mathematics has been undertaken in the Math project which was carried out under the bridge phase of the TIDE (Technology Initiative for Disabled and Elderly People) program [12]. The solutions of this project are based on speech output; mathematical expressions are read. This is a valuable alternative and a useful supplementing.

When we started our first project on computer-based aids for the blind in 1978, we wrote computer programs, which helped to integrate blind students into mainstream education [4, 5, 8]. Our programs taught blind children for instance to calculate with negative numbers and to use fractions. In theses programs we used the Marburg system. It was possible to write all mathematical characters with a computer, since we used the programming language APL. APL provided a character set of 256 characters in a time where the 7-bit-ASCII (American Standard Code for Information Interchange) with 128 characters was common. Each 6-dot-character, which appeared on the braille line was represented by an APL-character on the screen. However, children, who learn among the sighted, are taught by teachers who normally are not familiar with braille, especially not with a braille mathematical notation with several rules. Often more than one character is necessary to represent what the sighted show with just one symbol. This turns out to a big problem, if one wants to pass on the advantages of computer based learning and working. A simple example may already show this:

---

[1] Universität Stuttgart, Institut für Informatik, Breitwiesenstr. 20 - 22, D-70565 Stuttgart

23+34 in 6-dot-braille is written as , in SMSB as .
In print this would be represented by #bc +#cd 23 +34.

In SMSB the blank before plus may be omitted.

The late K. P. Schönherr had developed 8-dot-braille-modules, which were inserted into our first braille display, the BD64. It was connected with a microcomputer, the IBM 5100, and took into work in 1980. BASIC and APL were the programming languages which were provided on this computer. Thus, the meanwhile developed 8-dot-mathematics notation, the Stuttgarter Mathematikschrift für Blinde (SMSB) could be implemented and used. We oriented the SMSB on the Marburg system and discussed our proposal it with blind scientists, K. Britz, E. Hahn, H. Pinell, and D. Pograniczna in special workshops at Stuttgart.

Subsequently, we used SMSB in our computer program, which allowed blind students to discuss functions in analysis interactively. Students had for instance to find extreme points, or points of inflection. Moreover, we used SMSB, when we wrote a program to give blind students the possibility to differentiate elementary functions as rational functions, trigonometric functions or exponential and logarithm functions guided in a dialogue [8].

Though there exist many reasons for the use of a 6-dot braille, 8-dot braille reduces the gap between blind and sighted persons concerning the approach to written information. It is especially suited to give blind children, their parents and teachers an easier access to mathematics.

2 Concepts of the „Stuttgarter Mathematikschrift für Blinde“ SMSB

The „Stuttgarter Mathematikschrift für Blinde“ is an 8-dot-braille which includes small and capital letters, the numbers and mathematical symbols. SMSB provides three alphabets, the Latin, the old German and the Greek alphabet, to write letters which are used in mathematics. Capital letters differ from the small ones by the supplemented dot 7.

aA bB cC dD eE fF gG hH

The old German and the Greek letters are announced by  (dot 8) and  (dots 6, 8) respectively.

Each character of SMSB has a representation in braille and one in print. No SMSB-character stands for a string, a group of characters or even a word as it could be in contracted braille. That means, that each SMSB-character can be shown on a braille-output-device as well as on a screen and may be printed.

Numbers are written in the way which is in the meantime conventional in working with a computer. The letters from a to i are supplemented by dot 6, while zero is represented by the dots 3, 4 and 6. The number sign to announce numbers is not applied, it is not needed.

1 2 3 4 5 6 7 8 9 0

In the following we show mathematical strings in printed form and in braille. „SZ Schwa (Stuttgarter Zeichensatz *Schwarzschrift*) is the font, which we created to represent SMSB-ch equivalents on the screen and in printed form. „SZ Braille“ is our font to show the characters as are represented on braille-output-devices.

Meaning	*SZ Schwarzschrift*	*SZ Braille*
123+31=154	123+31=154	
15,678	15,678	
Roman number 25	XXV	

In the following, we describe symbols, show them in SZ Schwarzschrift and in SZ Braille and the tactile dots.

SMSB provides signs to indicate indices and just lowered characters by ↓ (dots 1, 6,

Exponents and raised characters or numbers are announced by ↑ (dots 4, 5, 7). Lowered

raised strings are announced by ↘ (dots 1, 5, 6, 8) and ↗ (dots 4, 5, 6, 7) respectiv

bars or e.g. arrows, which are put above or beneath characters or strings by ↖ (dots 4, 5, 8)

↙ (dots 4, 6, 8). In the same way one can underline strings or put brackets beneath them. E

announced string has to be terminated by ♥ (dots 1, 4, 5, 6, 7, 8).

Meaning	*SZ Schwarzschrift*	*SZ Braille*
X power 8	x↑8	
nth link of sequence a	a↓n	
n plus first link of a sequence a	a↘n+1♥	
Vector from A to B	↖→AB♥	

11. Symmetry

SMSB keeps symmetry of symbols as it can be shown with smaller and greater, < (dots 3, 5, 8) and > (dots 2, 6, 7) or with subset and superset, ⊂ (dots 1, 2, 6, 7) and ⊃ (dots 3, 5, 6, 8) respectively. Indicators for exponents and indices as shown above as ↑ (dots 4, 5, 7) and ↓ (dots 1, 6, 8) or ↗ (dots 4, 5, 6, 7) and ↘ (dots 1, 5, 6, 8).

The concept of symmetry is continued to brackets.

( ) [ ] { } ⟨ ⟩

12. Graphical Symbols

SMSB keeps graphic forms, which are used in writing and printing, as far as possible. So for instance the signs for smaller and greater as it was shown before. Other examples may be the arrows, hyphen and underscore, slash and backslash. All vertical lines are vertical columns of dots.

← → ↑ ↓ ↗ ↘ ↙ ↖ – _ / \ |

In the same way, we use symmetric forms for created graphical characters as for instance to represent fractions. In SMSB, a fraction starts with the opening bracket for fractions ▌ (dots 3, 4, 8). It is followed by the nominator, the line of the fraction ▄ (dots 2, 5, 7, 8), the denominator and the closing bracket of the fraction ▌ (dots 5, 6, 7).

Fractions may be also written in a simpler form as nominator / (dots 4, 7) denominator, where nominator and denominator may be written in brackets. The slash stands for the line.

Examples:

The fraction with the nominator x and the denominator x + t and the fraction with nominator 3a+2 and denominator 3c-1 are written as

The fraction with the nominator x and the denominator x + t and the fraction with nominator 3a+2 and denominator 3c-1 are written as

▌x▄x+t▌ ▌3a+2b▄3c-1▌ or as

x/(x+t) (3a+2b)/(3c-1)

The second notation may be easier for simple fractions as

11/2   5/4

2.3. Symbols in Arithmetic

The braille-characters for signs in arithmetic as + - * : start in the height of the dots 2 and 5. ... they can be identified by the reading fingers more easily.

\+   − + +   : ± ±   = ≠   < > < ≤ > ≥

2.4. Roots

In SMSB, a root starts with the symbol √ (dots 2, 3, 8) and ends with (dots 3,5,6,8) which also terminates strings. The power n of a root is written behind the root-symbol and announced as an exponent by ↑ (dots 4, 5, 7).

√t3   12xt4

A square root is just written like

√24x^16 -12

2.5. Examples

In this contribution, we conclude the description of SMSB with an arrangement of examples, wh... might be obvious or understandable at first glance. The printed version shows the representa... according to the Stuttgart fonts SZ Schwarzschrift and SZ Braille. It would appear on the screen ... printed here. The size of the characters can be chosen.

a or b	*4 power 3*	*Binomial coefficient 6 above 3*	*2 times square root of 3*
avb	4t3	6*3	2√3

^factions

f(x) =xt2 f"(x) =2x or df(x)//dx =2x

f(x,y) =2xy+3 5f/5y =2x

Égrals and Sums

iatb 3xt2 dx =[xt3]4atb =bt3 -at3

^d=1^ Tk(aii xti)

Itorices

M=(a^ik^> =<a^11^ a^12^ a^13^>
<a^21^ a^22^ a^23^>
<a^31^ a^32^ a^33^>

A matrix can also be written as a line, if the symbol ⇐ (dots 2, 8) to append rows is used:

M=(a^ik^> =<a^11^ a^12^ a^13^ #a^21^ a^22^ a^23^> ⇐a^31^ a^32^ a^33^>

Names of trigonometric functions and logarithms are written as in print:

sin cos arcsin arctan or arctg

log^10 10 =lg10 =1

10 log^e etx =10 lnetx =10x

The identification of special sets starts as in printed form with a vertical line, which is followed by the letter which identifies the set.

|N |G |Z |D |W   ∅ or { }

$|D = \{ x \in |R \; | \; x \neq 0 \}$

$0 \neg\in |N$ $0 \in |N \downarrow 0$

$|Z \vee \uparrow \neg \vee \downarrow 0 = |Z \setminus |N$

3. A Computer Based Working Environment to Write and to Print Mathematics

The components of a learning and working environment for blind students and scientists are a personal computer, a braille-output-device like a braille-display, a keyboard, a screen, a printer, and a braille-printer for embossed materials. Today, also a scanner and a pin-matrix-device for text and graphics may be reasonable.

Writing mathematics is efficient if the typed characters are represented in tactile form on a refreshable braille-output-device as a braille-line or a pin-matrix-device. The written text has to be embossed by a braille printer. An understandable printed version is also necessary to show the text in print to sighted persons. That means, fonts to represent the material on the screen and on a printer are needed. It should be possible to use a text processing system which is common or wide-spread.

3.1 Technical Preconditions

Using SMSB there exists one braille-symbol for each character, which is entered on the keyboard. It can be represented on a braille-display as well as on a braille-printer and on a connected laser-prints. This, however, is only possible if the used devices and the text-processing system provide the necessary character sets and fonts.

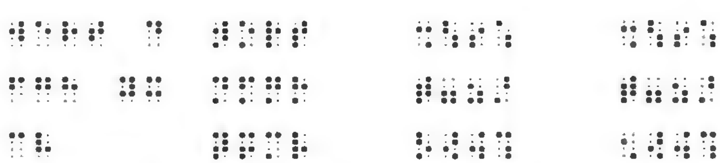
The braille-display must use the Stuttgart character set [11]. It consists of 256 characters and includes all characters of SMSB.

The keyboard should provide the SMSB-characters as it is the case in SEM (Stuttgarter *EMACS*), a DOS text processing system, which is based on the free and widely-spread EMACS. SEM has been developed at the University of Stuttgart in 1991. We continue to use its keyboard-layout if we write SMSB in Word for Windows, as we did in writing this paper.

The edition of the SMSB-description of 1998 is written with Word for Windows [10], while the edition of 1989 was written in LaTex [6].

We show twelve examples of keys of the SMSB-keyboard-layout. For each key we show the characters as they appear on the screen, if one just strikes the key, the key together with the key shift, together with the key Alt and together with the key Alt Gr. In each case we show the result in the fonts Times New Roman and in our created fonts „SZ Schwarzschrift“ (Stuttgarter Zeichensatz blackprint) and „SZ Braille.“ The representation in SZ Braille is also a picture of the output on the braille-output-device.

Font	*key*	*with shift*	*with Alt*	*with Alt Gr*
Times Roman	wert	WERT	ûî¶Ô	ûí¶Ô
	dfgh	DFGH	íÓöÕ	íÓöÕ
	yxcv	XYCV	É»ïÏ	É»ïÏ
SZ Schwarzschrift	wert	WERT	√ε→ i	√ε→ i
	dfgh	DFGH	∅∘≥↑	θ⊙≥T
	yxcv	YXCV	⊂⊃∩∪	⇐⊃∩∪
SZ Braille				



We applied mnemonics in overlaying the keys as far as possible. We oriented ourselves by the German language. We use for instance w for Wurzel (root), e for Epsilon (epsilon), r for rechts (right).

4 A Computer Based Working Environment to Read Mathematics

To write mathematical expressions, an adequate notation is necessary. It has to satisfy some requirements, if it should be used by blind and understood by sighted persons. The notations should be similar to the notation of the sighted in so far that symbols should be translated one by one from one notation into the other. This is especially true if an immediate translation from one notation to the other should be possible as it is urgent in schools where sighted and blind students learn together and where sighted teacher should understand at once, what a student writes. Only then the teacher and the parents are able to help. With such preconditions a computer based learning and working environment can be established.

It is more complicated to simulate the reading-process by a computer program. This is necessary if we want to translate printed or electronic material with mathematical texts into a representation for the blind. Until today we do not know OCR-software which interprets formulas. However, it is possible to show mathematical symbols and strings as a tactile picture [9]. If a text is written as a LaTex-document, the algebra and analysis expressions could be translated by a computer-program into a braille-notation, into SMSB, too. Graphics which are defined in LaTex have to be transformed into a tactile graphical version.

In the meantime, also in HTML, the language of the World Wide Web, there exists a mathematics notation. Since it is a linear representation, there is a good chance to translate mathematical expressions into an appropriate braille-representation.

The solution of these problems will be very helpful. They should not be confused with the need of writing mathematics at school, university or at work. The precondition for a reasonable working is an adequate notation, a notation which is used all over the world.

5. Conclusion and Outlook

We hope that blind children, students and scientists will be able to write mathematics in a reasonable notation if they use the advantages of modern technology. Until today, there exist still different opinions about 6- or 8-dot-braille. May be a convincing 8-dot notation will lead to an international solution.

6. Acknowledgement

First of all I want to thank our blind programmer Alfred Werner, who has been working in our research group AIB (Applied Informatics for the Blind) since 1978. He helped to create SMSB at the end of the seventies and in the beginning of the eighties. He wrote SEM (Stuttgart EMACS), a German version of the text processing system EMACS under DOS in 1991. SEM showed already to each character, which was entered on the keyboard, a character on the screen. These characters were designed and implemented by A. Werner together with Wolfgang A. Kreissl. They also proposed the layout of the SMSB-keyboard, which is essentially still in use, and provided the possibility to print out the written documents on a matrix-printer. Finally, Uwe Christian revised the Stuttgart character set and created the fonts for the laser printer. He made it possible to write SMSB under Word for Windows. I am very happy to work together with people who patiently help to continue this kind of work, which is never finished, because it is based on technical preconditions and facilities, which tend to change from time to time.

References

[1] Epheser, H., Pograniczna, D., Britz, K., Internationale Mathematikschrift für Blinde, in: Hertlein, J., Witte, R. F. V. (Hrsg.), Marburger Systematiken der Blindenschrift (Teil 6), Deutsche Blindenstudienanstalt, Marburg, 1992.

[2] Nemeth, A., The Nemeth code of braille and scientific notation, American Printing House For The Blind, Louisville, KY, 1972.

[3] Gardner, J., http://dots.physics.orst.edu/triangle.html.

[4] Schweikhardt, W., A Computer Based Education System for the Blind, in: Lavington, S. H. (ed.), Information Processing 80, pp 951-954, North Holland Publishing Company, 1980.

[5] Schweikhardt, W., The Impact of Micro-Computers on the Education of Blind Students, in: Lewis, R. and Tagg, D. (eds.), Computers in Education, North Holland Publishing Company, pp. 461-467, 1981.

[6] Schweikhardt, W., Stuttgarter Mathematikschrift für Blinde, Vorschlag für eine 8-Punkt-Mathematikschrift für Blinde, Universität Stuttgart, Institut für Informatik, 1983 and 1989.

[7] Schweikhardt, W., Rechnerunterstützte Übertragung einer mathematischen Formelsammlung in die Stuttgarter Mathematikschrift für Blinde, in: Ebersold, J. M., Schwyter, Th., Slaby, W. A. (Hrsg.), Computerised Braille Production, Proceedings of the 5th International Workshop, pp. 239-249, Winterthur, 1985.

[8] Schweikhardt, W., Rechnerunterstütztes Lernen für Blinde zur Förderung der Integration blinder Schüler in die Regelschule, in: LOG IN, Band 2, Heft 3, pp. 26-30, 1982.

[9] Schweikhardt, W., Interactive Exploring of Printed Documents by Blind People, in: Klaus, J. et al. (eds.), Interdisciplinary Aspects on Computers Helping People with Special Needs, 5th International Conference, ICCHP '96, pp. 451-458, Linz, Austria, July 1996.

[10] Schweikhardt, W., Stuttgarter Mathematikschrift für Blinde, Universität Stuttgart, Institut für Informatik, 1998.

[11] Schweikhardt, W., Christian, U., Werner, A., Neufassung des Stuttgarter Zeichensatzes mit allen Zeichen der Stuttgarter Mathematikschrift für Blinde, Technical Report, Universität Stuttgart, Institut für Informatik, August 1997.

[12] Weber, G. et al., Mathematical Access for Technology and Science for Visually Disabled People, TIDE MATHS Project Final Report, 1997.

Human-Computer Interaction (1)

-ADAPTIVE DIRECTNESS- A PRODUCT ORIENTED APPROACH TO EVALUATE SPECIFIC HCI SITUATIONS

Thomas Kahlisch, Wolfgang Wünschmann [1]

istract

...tive technology for people with special needs is very often combined with a need of additional ...t for managing Human Computer Interaction. Different strategies for handling the user ...face in such cases have to be compared for finding the best way towards high ergonomic ...ity. Standardized methods for usability testing have to be adapted to special needs of special ...s. For this purpose a new approach is described to define an objective product oriented ...ure, the Adaptive Directness. This measure has been used to evaluate a prototype of a table ...ter for blind users. Based on the related results a model has been developed for supporting the ...eralization or modification of the definition of Adaptive Directness. Hints at open problems and ...ntial next steps are given.

1. Introduction

...re are a lot of reasons for the need of testing the usability of technical products and services. ...ring to assistive technologies the most important reasons are:

- ...ility improvement in general
- ...omparison of different strategies
- analysis of week points in training programs

[1] ...den University of Technology, Institute for Information Systems, D 01062 Dresden
...one:+49 351 463 8467
...Mail: kahlisch@dzb.de; wuenschmann@inf.tu-dresden.de

Electronic documents have an increasing importance in all areas of social life (education, occupation, leisure). That is the reason that usability tests for technical aids helping people to get access to electronic documents have an extraordinary importance.

In contradiction to the demanding situation there are no in general accepted methods and models for a quantitative description of the influences of assistive technologies on the usability of electronic documents for people with special needs.

Based on the experience of the authors exist the following reasons for the difficulties of developing evaluation methods for specific HCI situations, using assisitive technology like Braille displays or synthetic speech:

I. The number of well trained and experienced blind computer users is small.
And assisitive technology is expansive.
That makes user oriented field testing methods difficult to use.

II. Assistive Technology causes complex dialog situations. An evaluation method, which can handle this complexity, needs to distinguish between the following aspects:
- problems belonging to the application (product),
- definition of the task, that has to be solved by the user,
- influences on the user constitution and
- the specific problems of using assisitive technology (e. g. limitation of an character based user interface).

This paper describes a formal product oriented approach to evaluate specific HCI situations, using assisitive technology.
The idea of this concept bases on the interaction room concept developed by Rauterberg [6].

The primary goal of the authors approach to define an objective measure for evaluating assistive technologies has been focused on contributions to improve the access to electronic documents; From this point of view HCI methods used by sighted persons and those used by blind persons have been analized in detail. The aim to find a concentrated expression to summarize the findings of the comparison has lead to the definition of Adaptive Directness.

tions from practical situations of blind students seem to justify the expectation measure could be used in very different dialog situations based on assisitive technologies.

Adaptive Directness

The current paradigm of assisitve Technology like screen readers gets the blind computer user in a ...ded by second order' interaction situation. The bridge software handles the information exchange between the user and the computer system.

...ase of the limitation of the tactile or speech oriented adaptive displays, the bridge software ...s to have flexible control mechanisms, which allow the blind user to explore the richness of the ...able information on the graphic user interface. In most cases the exploration process is ...ssary and costs plenty of time.

By using Rauterberg 'Interaction Point Model' it is possible, to quantify the number of elements of a given interaction context which are available to the sighted user. By using this method, it is also ...ble to quantify the number of elements which are available in the same situation to the blind

...erberg names the following elements of an interaction context:

Interaction points are elements of an user interface like mouse sensitive arrays, menu items or hot keys.

Objects are elements which can be changed by interactions (e.g. windows, forms or pictures).

To determine the relationship of the number of elements available to the sighted user to the number of elements available to the blind user the following expression can be calculated for each interaction context:

$$AR = \#IPF / \#IPD \quad (1)$$

AR adaptive relationship of interaction points for a given interaction context

#IPD number of interaction points in a given interaction context

#IPF number of interaction points available to the blind user in a given interaction context

The same relationship can be calculated for the objects in a given interaction context as well:

$$ARO = \#OF / \#OD \qquad (2)$$

ARO adaptive relationship of objects in a given interaction context

#OD number of objects in the given interaction context

#OF number of objects available to the blind user

To get full control over a dialog situation the blind user usually needs to explore a given interac context by using particular reading functions provided by the assisitive technology.
The expenditure of this exploration process causes in a product oriented approach in the following items:

- the sights of the adaptive display and
- the cleverness of exploration functions.

In an empiric study the following equation was found to calculate this expenditure:

$$EXP = C^{**}MS \qquad (3)$$

EXP expenditure for a blind computer user to get full control over a given interaction context

C empiric determined factor to rate the sights of the adaptive display and the cleverness of the exploration process

MS minimal number of exploration functions

By combining the Adaptive Relationship of interaction points and objects with the EXP function it is possible, to create an product oriented approach to quantify the difference between the information available for a sighted user to a blind user.
This measure is called 'Adaptive Directness'.
It can be calculated for a dialog system in the following way:

$$AD = SUM[d=1;n]\ ARd * EXPd \qquad (4)$$

The Adaptive Directness for the objects of an dialog system can be calculated as follows:

$$ADO = SUM[d=1;n]\ AROd * EXPd \qquad (5)$$

Usability evaluation of specific HCI situation

The relevance of the described approach was check out in a usability evaluation of a particular dialogue system developed at the Dresden University of Technology.

The hardware is a prototype of an special Table Browser for blind computer users. The program provides blind students access to complex tables in hypertext based study materials. The features and functionality of the browser is described at [5].

By using the described approach, other formal description methods [7] and special KLM [1] operators it is possible, to evaluate all components of usability (effectiveness, efficiency and user satisfaction) [4] for a particular dialog system.

The evaluation process of the Table Browser bases on solving three different example tasks. The steps which have to be made by the user and the belonging results are displayed in tables and other formal descriptions like interaction graphs. By using this formal product and task oriented description method, it was possible, to prove, that the evaluated software system provides blind computer users presentation and navigation functions to read complex tables in different ways.

The detailed evaluation was supported by using the following special KLM operators:

Rb-Operator: The time to read one Braille character from a Braille display (Rb=0.28)

Hb-Operator: The time to move the hand from the keyboard to the Braille display and backward (Hb=0.98).

The special operators have been detected in a self test by the author. The results are very close to earlier published values [2].

The results of the KLM have proved the following assumptions:

- The time for reading represents the highest expenditure for a blind computer user during an interaction situation.
- Well designed presentation and navigation functions are useful to reduce the reading expenditure.

Further research will show whether the described approach can be used for different specific HCI situations and how the product and task oriented model can be expanded with user oriented methods.

4. Model for generalization of Adaptive Directness

4.1 Proposal on a model structure

For further investigations on the usefullness of the measure Adaptive Directness a model as shown in figure 1 should be used.

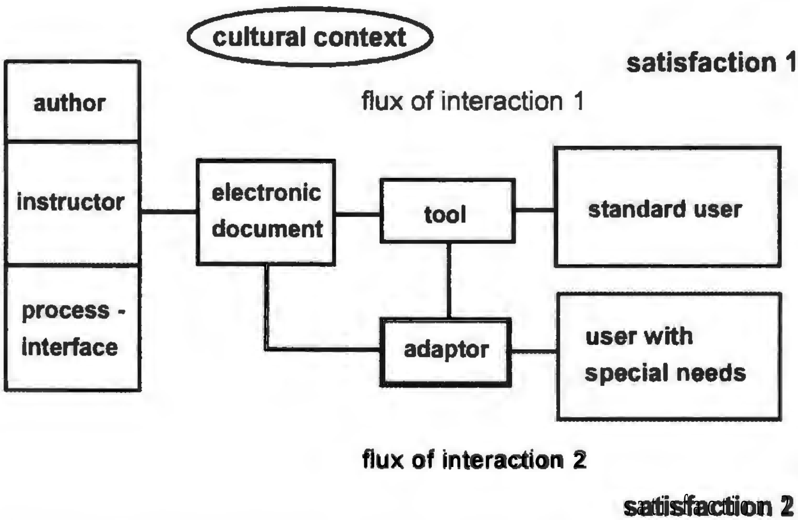


Figure 1: Model for generalization of the concept of „Adaptive Directness“

4.2 Importance of electronic documents

Figure 1 is closely related to the by the authors investigated situation of access to electronic documents. They can in general have very several complexity physical representation. They follow the tendency to become in the near future more and more a substitute of real existing objects. Electronic documents include both components with descriptive content and components with instructive content (description of tasks). Instructions are sometimes given implicitly or are based

Bnvetitions (e.g. reading a book). Consequently the electronic document is combined with an author (generator of descriptive content) and with an instructor (generator of tasks) in figure 1. A parallel situated process interface is an interface to general processes of production or to other activities substituting other authors and instructors.

Types of generators influence the level of interaction barriers and in this way the level of Adaptive Directness and in the end the usability of the technical system.

Transformation of interaction flux

Processing of documents needs always tools (combinations of hardware and software). The standard user (e.g. a sighted user) manipulates the document by managing the interaction flux I1, a flux of combined input-output activities integrated with planning and decision activities.

A user with special needs (user of assistive technologies) has access to the electronic document and to the tool by an adaptor (technical aid) under transformation of the interaction flux I1 into the interaction flux I2. Basis of both interaction fluxes are the interaction points #IPD, #OD, #IPF, #OF ... ting additional interaction steps for managing the transformation.

Implementation of the adapter in figure 1 with one path to the electronic document and another into the tool shall emphasize the possibility to integrate very different strategies of assisitive technologies. For instance technologies exist that activate the given document not directlybut by ... an on-line or off-line converted version of the document. Therefore additional effort is necessary what has to be taken into account for calculating Adaptive Directness.

Task and user based approach

The expressions of equations (4) and (5) include all accessible interaction points in an equivalent manner. With the introduction of the task depending interaction flux I1 a task related application of Adaptive Directnes has been prepared. The task reference can be carried out by a task depending vector VT (#IPO, #OD). This vector can be explained as a histrogram function.

Very often the time necessary for performing a task is used as a component of usability of the product under consideration. Taking into account a user specific vector of time periods to perform elementary interactions the performance of a complete task can be discussed task and user specific with reference to Adaptive Directness.

4.5 Satisfaction

The functional relation between satisfaction S of an user performing a task and the amount of additional interaction steps IS necessary to perform the same task by using assisitive technologic S = f(IS, AR, AO,..) is very complicated. The satisfaction can be analized in standard situation by using an extensive set of scaled criteria [4]. To decrease the expense for such evaluation method empiric functions should be introduced comparable with the approach of equation (3). But there is only negligible knowledge up to now on typical structures of such functions for typical applications of assisitve technologies. The cultural context of usage assisitve technologies will influence the empiric functions.

5. Perspective comments

The measure Adaptive Directness can be used to formulate statements on the quality of assisitvt technologies. The comparison of accessability of interaction points at the user interface of a tool for handling electronic documents with and without usage of assisitve technologies is the basic principle for the definition of Adaptive Directness.

The given prposal of a model for expanding the product oriented approach to a task and user oriented measure can help to expand the knowledge on usability of assistive technologies and can help to harmonize the methods for discussing evaluation results.

The need of comprimation details coming from usability tests to a statement on satisfaction of persons using assisitve technologies will furthermore exist as a serious problem of usability tests In connection to this problem a very hot topic with high relevance to Adaptive Directness exists in the following question: What are the driving forces for the tendncy that users of assistive technology sometimes change over from doing all necessary interaction steps for having a maximal interaction competence to a management method doing not all necessary interaction steps for full interaction competence and accepting an increased risk of interaction errors.

The challenge to use the measure Adaptive Directness during prototyping of new products belonging to the field of assistive technologies is another very important field for perspective investigations. Puplished guidelines on methods for software development with high ergonomic quality [3] can serve as fundamental knowledge but they have to be adapted to specific needs and rights concerning users of assisitve technologies.

6 References

[1] CARD, S.K.; MORAN, T.P.; NEWELL, A.: The Keystroke-Level Model for User Perfomance Time with Interactive Systems. In: Communications of the ACM 23. 1980. S. 327-356.

[2] DENNINGHAUS, E.: Die Foerderung der Lesegeschwindigkeit bei blinden und sehbehinderten Jugendlichen und jungen Erwachsenen. In: Blind-Sehbehindert. Nr. 2 96. S. 95-100.

[3] DZIDA, W.: International User-Interface Standardization. -In: Allen B. Tucker, Jr. (Ed.): The computer science and engineering handbook. -Florida: CRC Press. 1997. pp. 1474 - 1493

[4] ISO 9241 Ergonomic requirements for office work with visual display terminals (VDTs)-Part 11: Guidance on usability (Version ISO/DIS 9241-11.2 1997)

[5] KAHLISCH, TH.: Making hypertext based study material accessible and user friendly to blind students. Sixth International World Wide Web conference. Accessibility Track. Santa Clara 6. April 1997.

[6] RAUTERBERG, M.: Ein Konzept zur Quantifizierung software-ergonomischer Richtlinien. Zürich 1995. ISBN 3-906509-11-7.

[7] WÜNSCHMANN, W.: Grundlagen der Mensch-Maschine-Kommunikation. Vorlesungsskript. TU Dresden. Fakultaet Informatik. Sommersemester 1997.

Bio-Signaling for Control of Graphic User Interfaces

Dr. Heide-Rose Vatterrott, Sybille Hambach,

Oliver Hein, Thomas Heider

Conventional input devices are not convenient for people with special physical conditions. To provide new means of controlling complex interactive systems, frequent biological signals can be used. This paper explains technical devices for registering and analyzing biological signals and describes requirements to be met by hardware and software in order to satisfy medical demands. A distinction is drawn between the control of newly developed special applications and the direct or indirect control of conventional software programs. Different devices and applications are described. All of them have proven to be suitable for determining and training the perception and knowledge of patients undergoing rehabilitation as well as enabling these patients to communicate in a self-determined manner.

1. Motivation

To use a personal computer people need to learn and perform special tasks. They must be able to handle input devices designed for everyday work and normal physical conditions. However, conventional input devices are not convenient for people carrying out exceptional tasks. Likewise, they cannot be used by people with special physical conditions, e.g. the elderly or disabled. To give these people the chance to use a computer for work or everyday activities it is necessary to provide them with special input devices.

One common way of developing new input devices is to modify conventional tools. Another method is based on bio-signaling. Every person transmits frequent biological signals. Some of them, e.g. eye movement or muscle contraction, may be controlled consciously without training.

For this reason, computer software utilizing biological signals can be used by people performing exceptional work tasks or with specific physical conditions.

The BioCom project [1] described in this paper is being carried out in close cooperation with specialists at the Fachkrankenhaus Waldeck, a hospital for early rehabilitation. Its aim is to develop interaction techniques for totally disabled persons, who cannot interact verbally as a result of [illegible] or neurological diseases. Such persons are motorically restricted and thus not able to use conventional input devices. However, they do understand speech and can express themselves by means of language.

This paper explains technical devices for registering and analyzing biological signals that have been used within the context of BioCom. Also described are requirements to be met by hardware and software in order to satisfy medical demands and the applications developed to utilize biological signals as an input device for totally disabled people and people undergoing rehabilitation.

All BioCom applications have been tested extensively in the Fachkrankenhaus Waldeck. They have proven to be suitable for determining and training the perception and knowledge of patients undergoing rehabilitation as well as enabling these patients to communicate in a self-determined manner.

2 BIOLOGICAL SIGNALS

Biological signals originate from the human body and can be measured as or transformed into voltages. In order to be used to control a computer they have to meet two important preconditions:

- they need to be controlled consciously without training
- they need to be measured easily and non-intrusively

2.1 Suitable Signals

Eye movement, muscle contraction and speech comply with the preconditions for biological signals capable of being utilized to control a computer. As the BioCom project aims at developing interaction techniques for persons not able to interact verbally, speech cannot be used. This leaves eye movement and muscle contraction as the only suitable signals.

2.1.1. Eye movement

Eye movements – up, down, left, right – can be controlled consciously without training. They may be used to control the input focus of interactive systems. The movements are easy to track using optical as well as bio-electrical gauges (see next subsection). The permanent saccadic movement of the pupil and blinking have to be filtered.

2.1.2. Muscle contraction

Muscle contractions – tensed, relaxed – can also be controlled consciously without training. They may be used to trigger actions in interactive systems. Contractions have to be recorded by bio-electrical gauges. The permanent unconscious tension of muscles can be filtered easily.

2.2. Devices for Registering Biological Signals

Devices for registering biological signals are extensively used in the field of medicine. Some well-known examples are means of recording heartbeat (ECG), muscle contraction (EMG) and eye movement (EOG). This technology has been perfected over the past few years. Systems have been used to record and analyze biological signals for medical diagnosis as well as to control prostheses. Current research is studying the utilization of biological signals for controlling complex interactive systems [1, 2, 3].

The optical and bio-electrical gauges used in the context of BioCom are based on well-known methods for registering biological signals [1, 3, 4] and have been utilized in the field of medicine for some years. Of all the different methods and gauges tested, these two have been the most promising with regard to usability in this new area of application.

1.1.1. Optical gauges

HCom utilizes the Remote Eye-tracking System from SensoMotoric Instruments GmbH Teltow (Germany) [5] based on the pupil center / corneal reflection (PCCR) technique [4]. As totally disabled persons are often not able to move their head, the stationary camera used in this system does not represent any restriction. The PCCR technique is the most promising from a technical point of view. However, the gauges are still too expensive to be used extensively.

1.1.2. Bio-electrical gauges

HCom utilizes the Computer Nystagmograph Analyzer from Hortmann Neuro-Otometrie GmbH Menzlingen (Germany) [6] for eye tracking and detecting muscle contractions. With this system, correct placement of sensor pads and the cable connection of pads and gauges is easy and reliable. The technical equipment is not as expensive as optical gauges and is paid by medical insurance companies.

1.3. Analysis of Significant Courses of Signals

After biological signals have been recorded by means of optical or bio-electrical gauges, the signals have to be analyzed for significance. The permanent saccadic movement of the pupil and the permanent unconscious tension of muscles have to be filtered to detect significant courses of signals. These are different for every person and each session. For this reason, the gauges have to be calibrated before using the recorded signals for the control of interactive systems. They may have to be re-calibrated during a session in order to compensate for signal variation caused by altered concentration or perspiration.

By calibrating the gauges and adapting various parameters, it is possible to adjust the measuring devices to suit every single person for each session. This is most important for all technical devices to be used in the field of rehabilitation because of the wide range of impairments and the individuality of each single patient. Following calibration of the gauges, the biological signals, e.g. eye movement and muscle contraction, can be used to control complex user interfaces.

3. MEDICAL DEMANDS

The software applications developed in association with BioCom are intended to support people impaired by stroke, injuries or from birth in

- therapy, e.g. in determining and training perception and knowledge, and
- self-determined communication, e.g. in expressing everyday needs.

These people are paralyzed and not able to communicate verbally. Eye movement, winking and muscle contractions are their only means of communication and sometimes their only movement of any nature.

The following requirements are partly known from the literature [7] and partly derived from testing the first prototype applications of BioCom at the Fachkrankenhaus Waldeck. These tests have shown that all the following requirements are realistic as well as relevant.

3.1. General Requirements

General requirements to be met by hardware and software include the following:

- easy to operate and user-friendly for therapists and patients (the therapist needs to be able to concentrate on the patient, not on the technology),
- use of alternative input devices dependent on the patient (some patients are able to move muscles, others may even have limited eye movement etc.),
- tolerant of errors (patients have no mental ability to handle complicated error messages),
- no or minimal additional strain for the patient (impairment and treatment are a burden, requiring all physical and mental strength),
- affordable (all equipment has to be payable by medical insurance companies).

3.2. Hardware Requirements

The hardware to be used to record and analyze biological signals as well as the computer itself should meet the following requirements:

- all technology has to comply with statutory safety regulations.

- optical and bio-electrical gauges have to be robust and flexible in their use as well as easy to handle and to calibrate (requirements for gauges used at home are much higher than for those used in medical diagnosis only),
- the computer monitor has to have sufficiently high resolution and a clear image (patients have to "stare" at the monitor if the input focus is controlled by eye movement; some patients may also have poor eyesight or visual defects).

B. Software Requirements

All applications controlled by means of biological signals have to give acoustic and / or visual feedback. In addition, applications used for therapy should meet the following requirements:

- clear formulation of tasks and mediation of a sense of achievement,
- increased demands and random generation of new tasks,
- change of media (text, images, sound) and speed control,
- results should be visible for patients as well as therapists.

Additionally, arrangements should be made for some kind of authoring tool to make it easy to provide new tasks with a similar structure, e.g. new mathematics exercises.

4. APPLICATIONS

The new input devices for biological signals can be used to control complex interactive systems. A distinction can be drawn between the control of newly developed special applications and the direct or indirect control of conventional applications.

4.1. Development of Special Applications

Applications with a specially developed user interface can be adapted perfectly to the new input devices. By developing completely new applications, it is possible to adapt the interaction technique to the symptoms of a group of patients or even a single patient. It is even imaginable to develop software for patients with eye movement in three or only two directions. On the other hand, such applications are highly specialized and can only be used by a very small group of

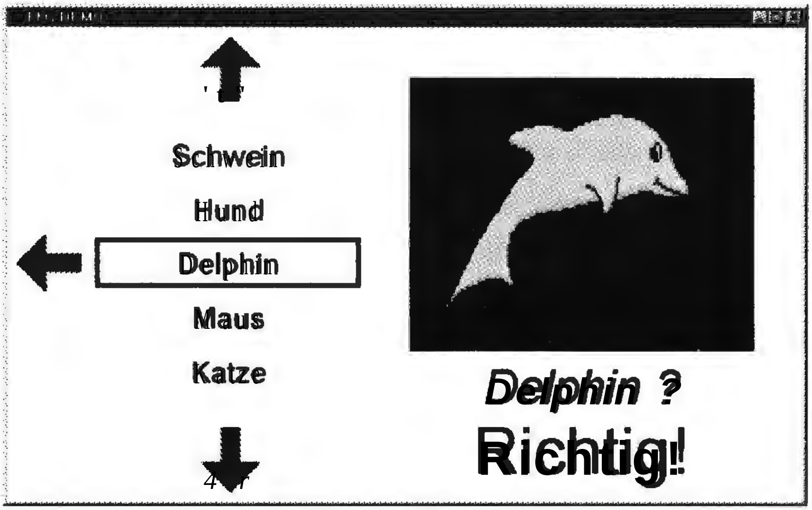


Figure 1: Naming of images

people. This disadvantage can be compensated for by providing some kind of authoring tool. Using this tool, the therapist can provide new questions / answers or set new exercises.

Specially developed applications are suitable for early rehabilitation. They are easy to use and can support training as well as therapy. It is best to provide a pool of different applications to be controlled by gauges for different biological signals in order to meet the individual needs of each patient.

One of the applications in this category is a program for naming images. It was developed to test the perception and knowledge of a child disabled from birth and meets all the demands established in the previous chapter. The graphic interface shown in Figure 1 is used with a bio-electrical gauge. All the tasks, e.g. images and possible names, are specified in a text file. The program reads the file, displays the first image with corresponding names, waits for input and than shows the second image with corresponding names etc. The input focus, e.g. the highlighting of a name, is controlled by looking up or down. A word is selected by looking briefly to the left (Selection by means of muscle contraction is possible, too). After selecting the correct name, the next image appears. The child the program was developed for enjoyed naming the images. This was the first time the boy could show what he had learned in school.

4.1. Control of Conventional Applications

Conventional applications controlled using standard input devices, e.g. keyboard and mouse, have been developed for every-day tasks. These applications can be controlled by the new input devices either directly or indirectly.

4.1.1. Direct Control

The software for analyzing biological signals is used in the same way as a simple software driver and directly connected to the software application. One example of this technique is the control of the cursor of a computer game by means of eye movements and the triggering of actions through muscle contractions.

Using this technique makes it possible to control conventional applications. However, the problems associated with conventional software, e.g. complexity or long training periods, remain. In addition, conventional software is suited only to the existing hardware. For this reason, the use of new input devices either requires more steps to complete a task or renders some of the functions entirely useless.

It can be concluded that the direct control of graphic user interfaces by means of bio-signaling is not suitable for early rehabilitation. The reasons for this include the complexity of conventional applications and the time required to master the technology. However, applications in this category may be adequate for use in later stages of rehabilitation.

4.1.2. Indirect Control

The software for analyzing biological signals is connected to standard applications via a software program with a special user interface. One example of this category is a virtual keyboard to control a standard text editor.

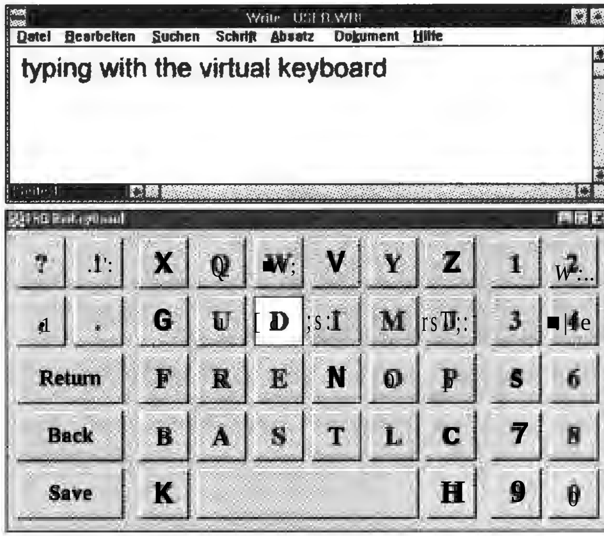


Figure 2: Virtual keyboard as input device for a standard editor

Indirect control allows for the user interface to be adapted to bio-signaling as an input device. Thi makes the software easier to learn and operate. However, it is probably not possible to access all the conventional application's functions.

The indirect control of conventional applications through software programs with an adapted user interface is suitable for self-determined communication. As patients are eager to express them selves, they can easily control the adapted application. The reduction of functionality of the standard application does not present a problem as most of the functions are not essential for communication purposes, e.g. change of fonts or correct spelling in a word processor.

Different virtual keyboards have been developed within the context of BioCom. The keyboard shown in Figure 2 is controlled by means of a bio-electrical gauge. The letter currently focused is highlighted (letter D in Figure 2) and than selected by being looked at for the time set. Change to neighboring letters is carried out by looking to the left, to the right, up or down. Letters are arranged according to their frequency in the German language to facilitate fast typing even for non proficient users. The keyboard is large enough to be visible for patients sitting in beds or wheelchairs.

5. FUTURE WORK

Future work includes the development of additional applications, e.g. more software programs to be used in therapy and training, as well as the utilization of new devices for registering biological signals, e.g. speech or sound recognition.

A joint interface is under development for new and indirectly controlled applications on the one hand and devices for recording biological signals on the other. This interface will make it easier to interchange input devices dependent on the patient and his symptoms.

At the same time, therapists and patients at the Fachkrankenhaus Waldeck are using the new interaction techniques in therapy and training. Their help and advice has been invaluable in the past and will be vital for all future development.

6 REFERENCES

[1] VATTERROTT, H.-R. BioKom. *Computer Graphik Topics 6196*, Vol.8, 28-29.

[2] LUSTED, H. S., and KNAPP R. B. Computersteuerung mit Nervenimpulsen. *Spektrum der Wissenschaft*, (December 1996), 72-78.

[3] GIPS, J., and OLIVIERY, P. Eagle Eyes: An Eye Control System for Persons with Disabilities. *The Eleventh International Conference on Technology and Persons with Disabilities* (Los Angeles, March 1996). Available as http://www.cs.bc.edu/~gips/EagleEyes/EEReport.html.

[4] Eye Tracking. Available as http://www.asel.udel.edu/ ~joyce/EyeTracking.html.

[5] SMI SensoMotoric Instruments GmbH, Technology Terrain Teltow, Potsdamer Str. 18a, D-14513 Teltow, Germany.

[6] Hortmann GmbH Neuro-Otometrie, Robert-Bosch-Str. 6, D-72654 Neckartenzlingen, Germany.

[7] GOBIET, W. Bedeutung der computergestützten Therapie in der neurotraumatologischen Rehabilitation. *Tagungsband Forschung und Praxis der Neurologischen Rehabilitation – 10 Jahre KURATÓRIUM ZNS* (Hamburg 1995), 285-294.

VIRTUAL SHAPE-MAKING SYSTEM FOR TACTILE RECOGNITION OF THE VISUALLY CHALLENGED

Yoshihiro IMAZEKI[1)]
Yoshimichi YONEZAWA[1)]
Kazunori ITOH[1)]

Abstract

The virtual shape-making system was made to help the visually challenged. The system consists of a personal computer, mechanical movements on two dimensional (2D) plane of X-Y recorder and force sensors mounted on a stick. A vertical stick on a pen-head of the recorder can move by force of fingers to any directions in an area where no shape is displayed, but can't move into the inner area off a displayed shape. Then the stick can move along the edge of the shape to recognize an image on the 2D-plane. The recognition rate was compared between shapes virtually displayed and real shapes made of cardboard in equal size by the same sensing procedure. It was shown to be same as each shape. The visually challenged can know many kinds of virtual shapes on the 2D-plane that are programmable and are stored in the memory of the personal computer. The system can be applied for making educational tool such as tactile letters, miniatures of big animals that are not perceptible by fingers, patterns for geometry, maps and scenographs.

1. Introduction

Yoshii et al. [4] reported that sensing point density of fingers on palm side was a hundred per square cm. When we touch an object by fingers, the contact area of fingers on it is 0.5 cm^2 per finger in average. It means that 250 sensing points are on five fingers contacting. The visually challenged get information by these sensing points. Generally they can't get all information of a large object at a time, though they can understand the shape and the size if an object is small. To get some more information, they slide their fingers along the edge of an object like a scanning sensor. They have more data for tactile recognition by a finger sliding than when the fingers are not moving. To make the system simpler, we intend to use the only data through a stick instead of sliding the first finger.

---

1) Faculty of Engineering, Shinshu University, Nagano 380-8553, Japan

1 Design of the system

1.1 Requirement for *virtual shape-making system*

Vitual tactile recognition systems for the visually challenged have been researched [1,2,3]. But the systems are too large and complex to be put into practice and deal with many channel data. We build up a *virtual shape-making system* that is simple and uses one channel consecutive data.

To recognize a shape through a stick, the visually challenged need to feel reaction to their fingers from the stick.

The functions of the system are:

1) The force sensors pick up a direction and magnitude of force applied by fingers, and transfer the information to the system.

2) A stick with force sensor can move freely on the plane and users feel reaction through the stick at the edge of a shape, so that the stick can't break in the shape as a forbidden area on 2D-plane.

3) The system has various kinds of shapes in computer memory and any users can know much kind of shapes at any time immediately.

1.2 Interface of the system

A couple of sensors for X-axis and Y-axis components of the force is necessary for the system. We use two strain gages (Kyowadengyo; KFG-IN-120-C1-11; the size of strain gage is 6mm^2) as sensors on a stick. The size of the stick is 6mm in diameter. So that it is easy to attach sensors to the stick and to handle it. The stick is made from acrylic resin. The strain gages as sensors are fixed around the stick for detecting X and Y components. Then the gages are making 90 degrees angle. The two output voltages of gages generate as a function of applied force on the stick respectively.

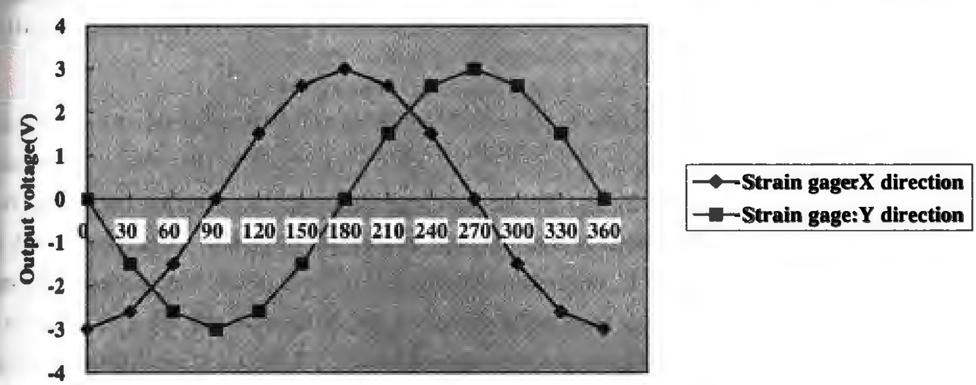


Fig.1 Output voltages of X and Y sensors as a function of applied force.

Fig.1 shows two output voltages of gages with 0.5kg applied on the stick as a function of force directions. The curve can be approximate to sine and cosine function respectively. From voltage data gained from X and Y sensors, the direction of force applied is easily calculated.

2.3. Mechanism of the system

We used an X-Y pen-recorder (Graphtec; WX2400) as mechanical 2D-movement. Fig.2 shows the diagram of the *virtual shape-making system*. A vertical stick with force sensors was placed on the pen-head. As a general characteristic of an X-Y recorder, the position of the pen-head is controlled by a couple of input voltages V_x and V_y. And a spot on the plane of an X-Y recorder corresponds to a couple of V_x and V_y, further an area (shapes) corresponds to groups of V_x and V_y.

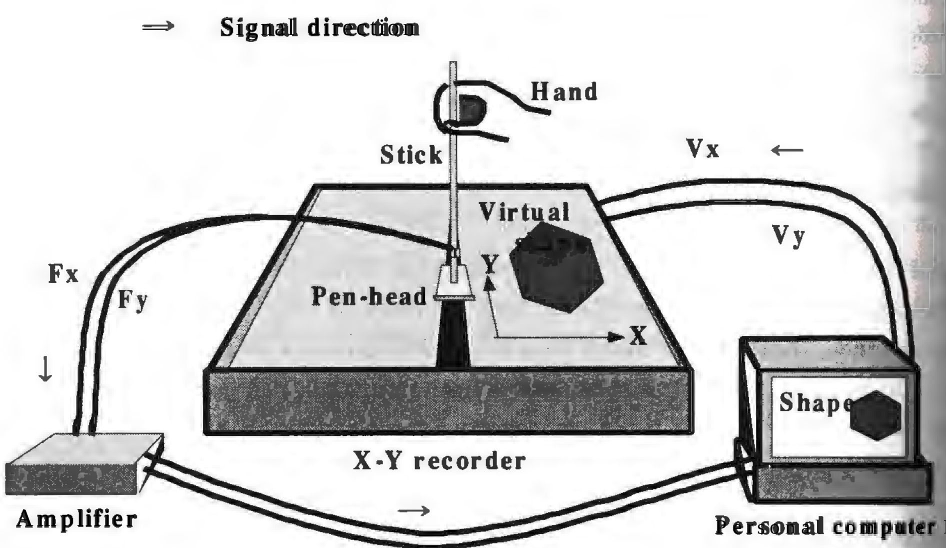


Fig.2 A diagram of the virtual system.

In the system, the personal computer continuously acquires the voltage signals from force sensors, F_x and F_y. When no force is applied on the stick, either of them shows 0 volt. But once the user adds force on the stick to move it, F_x and F_y change depending on the direction and strength of the applied force. The computer calculate the direction which the user want to move the stick to, and check whether there is an obstacle, the forbidden area just in front of the way to move. If there is no forbidden area, the stick moves a step to the direction and if there is, the stick can't move. In the latter condition, the user must change the direction of force until the stick is allowed to go ahead. It means the stick is on the new position after a small step movement and also that a couple of voltages V_x and V_y which control the position of the stick, are renewed respectively corresponding to the small one step. Thus, the stick travels on the 2D-plane by the repetition of the small steps, Fig.

ows the linearity between the direction of force on the stick and the direction calculated from pll voltages of force sensors on the stick. The forbidden area can be any pattern and size by the Inputer and the user can recognize the area as non-intruding one by the system

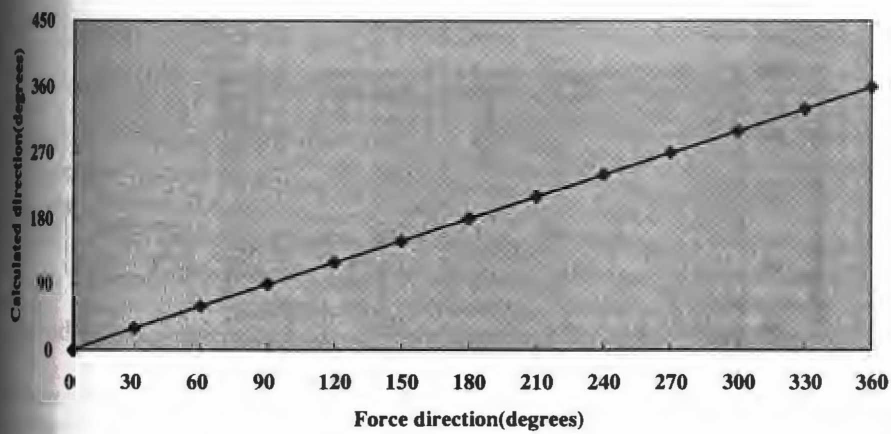


Fig.3 The linearity between force directions and calculated directions.

%4 shows a model of trace of stick, which meets a hexagon displayed on 2D-plane as non-iIrading area and walks around it.

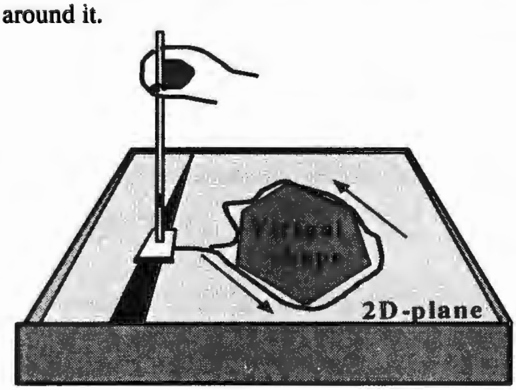


Fig.4 Model of trace of stick .

3 Recognition test

Un1D-plane, *the system* can display the patterns in which the stick can't intrude. The accuracy of ^Dgnition of the patterns was compared between shapes made on *the system* and real shapes. The ipes on *the system* were virtual objects on 2D-plane and the shapes of cardboard on the desk were ks(ones. The people with eye-bandage were asked to track both the shapes on *the system* and on

the desk by the stick in same size. The shapes we used were eight kinds of polygons (from tria
to decagon) with maximum width of 8cm. Fig.5 shows the recognition rate versus the numb
corners from triangle (N=3) to decagon (N=10). The accuracy decreases with increasing numb
corners, and is slightly lower for virtual shapes. But, it is surprising that the shape of hexag
recognized with 80% accuracy rate on the virtual system.

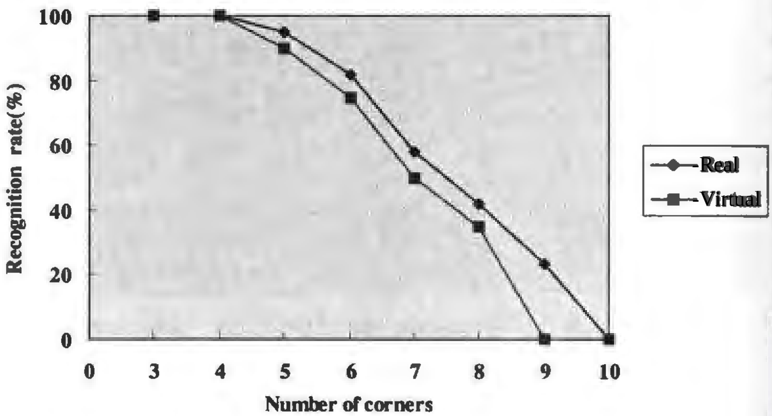


Fig.5 Recognition rate of the number of corners.

4. Conclusion

Using mechanical functions of an X-Y recorder, personal computer and force sensing stick,
made a new style *virtual shape-making system*. The recognition rate was compared between sha
made on the system and real ones, and was found to be almost the same. *The system* is useful in t
field of education of patterns, maps and scenographs for the visually challenged, and also in t
field of general use.

5. References

[1] IWATA, H., Institute of Engineering Mechanical movements, 8th Symposium on Human Interface, 1992.

[2] IWATA, H., Artificial Reality with Force-feedback, Development of Desktop Virtual Space with Compact Manipulator, Computer Graphics Vol.24, No.4, ACM SIGGRAPH, 1990.

[3] YOKOYAMA, A., ISHII, K., MAKINO, H., NAKASHIZU, M., A Simple VR System for Precise Opera Virtual Objects, Symposium on IEIEC, 1994.

[4] YOSHII, N., The present physiology, Nagai publishing bookstore, 1974.

ALTERNATIVE INPUT DEVICE FOR CONTROLLING THE COMPUTER WITH PERSONS CAPACITY

Vidas Lauruska[1]

Abstract

The paper describes an alternative input device, which could be used by a physically impaired person to control a computer or other electronic device. The device uses body capacitance for the control mechanism. The rationale for the design and the results of preliminary experiments are presented. The device is expected to undergo further refinement.

1. Introduction

The degree and ease of the computer access for severely physically impaired persons is limited, to a large degree, by the amount of their residual motor control. It is important to select an input device that makes maximum use of these capabilities. A person with only one or two reliably controllable movements can choose from a wide range of single or dual input devices [1].

The largest group of switches used as input devices is that of the devices designed to be activated by a person's head or limbs. Included in this group are simple lever or pressure switches, which can be operated by the movement of almost any part of the body, grip and thumb switches, pneumatic pressure switches, tilt and posture switches, sound activated switches and wobble sticks.

More sophisticated direct-access devices, such as mice, trackballs, joysticks, and touch-sensitive screens and boards, also used by impaired persons, present specific sets of problems. Computer mice require moderate to good hand and eye coordination, in addition to very fine motor control. Trackballs can alleviate some of these problems. Touch screens and boards are not accurate enough for finely detailed work, and are more suitable for use with software utilizing simple screen layouts.

The present paper describes an alternative input device utilizing person's body capacitance. The following requirements for such a device were identified as it should be simple to operate, and controllable through a single sensor, and support simple communication and control modes. The software used in the project are those used in another Siauliai University project [2].

2. Design

The input device allows impaired persons to control communication devices (writing communication) and environmental control systems (operating lights, a television set and etc.). A block diagram of the input device is presented in Fig. 1.

1 Faculty of technology, Siauliai university; 141 Vilniaus Str., Siauliai, LTU-5400, Lithuania, e-mail: vidas.lauruska@tsf.ktu.lt

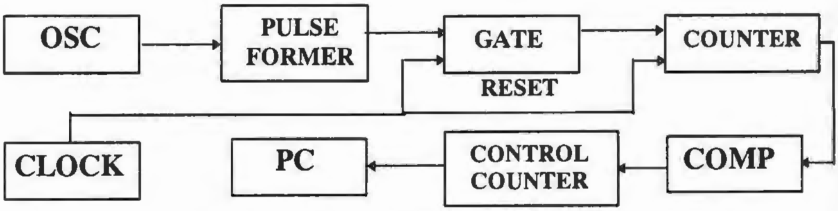


Fig.1. Block diagram of device

The input device consists of a high-frequency oscillator, gate, counter, comparator, control counter, and a clock for synchronization. The frequency of the oscillator is controlled by the varying body capacitance of the user. The sinusoidal output of the oscillator is converted into a pulse output by the pulse shaper, and applied to the gate circuit. Each clock pulse resets the counter, and allows the oscillator pulses to pass through the gate circuit. The comparator circuit outputs one pulse for every 53 input pulses from the counter. The control counter outputs one pulse every fourth comparator pulse to the computer. The control counter is designed to reduce the incidence of false signals due to involuntary movements of the user. The capacitive input sensor consists of a 4 cm² copper covered insulated plate. Movements of the user's body vary the capacitance of the plate, which is connected into the oscillator circuit of the device.

Tests were conducted to study the effects of the user on the frequency of the oscillator. The frequency of the oscillator was controlled by the distance between the user's hand and the sensor. Fig. 2 shows the test results. The frequency of the oscillator changed from its resting frequency by

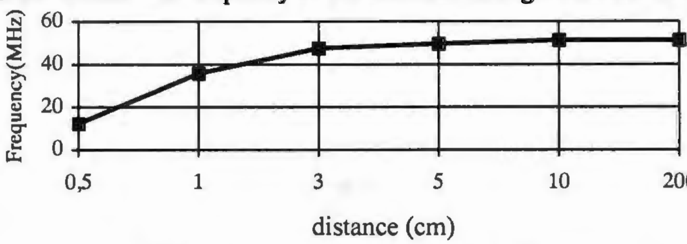


Fig.2. Frequency change as a function of distance

30% when the distance between the user's hand and the sensor was 1 cm and 80% when the distance was decreased to .5 cm. Similar results were obtained when investigating the control capability of the user's foot, finger, and other body parts.

3. Conclusion

The described device can be used as an alternative input device for persons with minimal reliable controlled physical movements. The advantages are the simplicity of the device, and zero requirements for physical strength to operate the device.

References

[1] COLVEN D. and DETHERIDGE T. A Common Terminology for Switch Controlled Software. A Guide for Software Developers, Published by ACE Centre, 1990, 37p.

[2] LAURUSKA, V., Computerized Control and Communication System for Handicapped, Proceedings of ICCHP, p.p. 515-519, 1996 (system which allows impaired persons to communicate using an eye closure system)

Session VIII

Access to Graphical User Interfaces

PROGRAMMING AND VISUALLY IMPAIRED PEOPLE

Ivan Kopeček and Adriana Jergová
Faculty of Informatics
Masaryk University
Botanická 68a, 602 00 Brno, Czech Republic
e-mail: kopecek@fi.muni.cz, ada@fi.muni.cz
http://www.fi.muni.cz/~kopecek/
http://www.fi.muni.cz/~jergova/

Abstract

Some problems of visually impaired programmers are discussed in the paper. The project of the dialog generator of program source code for visually impaired people that is developed at the Faculty of Informatics, Masaryk University is described. This system is supported by the library of procedures oriented to visually impaired and by blind user oriented hypertext system AUDIS. The possible use of declarative languages is also briefly discussed.

1. Introduction

Programming is a challenge as well as a new chance to break communication barrier and to realize activity for many visually impaired people. Although serious impairment of sight is a great limitation, new information technologies, especially speech synthesis and recognition, offer new possibilities how to support programming of visually impaired.

In the present paper we discuss some problems that the visually impaired meet when programming and learning programming and propose some ideas how to help them. Let us, however, point out that we make no aspirations to discuss all aspects of this very complex problem.

The basic problem blind programmers face to is the integrated environment that is absolutely inconvenient for them. Speech synthesizer and "screen access" ("screen reader") software utilities are used by blind users to obtain information that appears somewhere on the screen, depending situation. Even syntactical debugging can be a kind of nightmare for the sight impaired programmer (often as a consequence of imported errors).

Secondly, run-time libraries and utilities are not adapted to the sight impaired programmers. It is also very difficult to obtain an overview of possibilities offered by run-time libraries and to get information about the correct use of the chosen procedure or function.

2. Programming in Procedural Languages

Programming in classical procedural languages like C++ or Pascal is universal, popular and widespread. It is obvious that proposals to develop special integrated environment of these languages for blind users would be a bit platonic because only big software producers can realize! Nevertheless, simpler possibilities can be found how to partially solve the above mentioned problems of blind people programming. In this section we mention the project developed at the Faculty of Informatics, Masaryk University which is motivated by the aim to make programming easier for visually impaired students. The project consist of the following parts:

- Dialog generator of program source – special phonetically based editor of the program source that can be used to facilitate building of source code and to prevent syntactical bugs.
- Library of procedures for programming of visually impaired people.
- Application of the blind user oriented hypertext system AUDIS to provide a run-time library overview and to provide the information about the correct use of procedures and functions.

2.1 Dialog Generator of Source Code

The ideas behind the dialog generator of source code can be formulated as follows:

- Control of the system is supported by speech synthesis and recognition.
- Sound information is preferred whenever it is effective. It can be produced in the form of speech input and output, user speech comments and remarks to the source text, audio glances and earcons ([7]), environmental sounds ([5]) etc.
- Source code generation is performed in the form of man-machine dialog fully supported by speech synthesis and recognition.

- The dialog is performed in such a way that it prevents the users from committing syntactical (and to some extents semantical) errors and makes the building of the program simpler.
- The system provides the user with complex information about the actual position in the source code also related to the relevant program blocks.
- The system provides convenient reading (in audio form) and effective editing of program parts.

The following example illustrates the possible form of the dialog:

1 U: *assign*
2 S: [*] *left side*
3 U: sum
4 S: [*] sum *variable not declared; declare it?*
5 U: *yes*
6 S: [*] *type of the variable*
7 U: *integer*
8 S: [*] *right side*
9 U: *zero*
10 S: [*] *okay*
11 U: *read command*
12 S: [*] *assign zero to* sum
13 U: *standard cycle*
14 S: [*] *comment the cycle*
15 U: *the cycle sums numbers from one to one thousand*
16 S: [*] *enter the cycle variable*
17 U: i
18 S: [*] *the variable is not declared; declare it as an integer?*
19 U: *yes*
20 S: [*] *lower bound for the cycle variable*
21 U: 1
22 S: [*] *upper bound for the cycle variable*
23 U: 1000
24 S: [*] *define the body of the cycle; use end command to terminate the body*
25 U: *assign*
26 S: [*] [-] *left side*
27 U: *sm*
28 S: [*] [-] *variable* sm *not declared; declare it?*
29 U: *no*
30 S: [*] [-] *redefine the left side*
31 U: sum

32 S: [*] [-] ***define the right side***

33 U: sum+i;

34 S: [*] [-] *okay*

35 U: ***position***

36 S: [*] [-] ***you are inside the cycle the cycle sums numbers from one to one thousand***

37 U: ***end***

38 S: [*] ***okay***

The *italic* font is used to denote the spoken (synthesized) parts of the dialog (bold italics is used for different type the speech output). The underlined font stands for the keyboard typed information. The **bold** font means spelling. Special symbols [*] and [-] have the following meaning:

[*] – sound signal "your command was accepted;"

[-] – short beep informing that we are inside one program block (position inside two program blocks is signalized by two beeps, great number of beeps is expressed by combination of beeps of different duration).

The enumeration of the dialog items relates them to the following comments:

11. The user asks the system to read the generated command.

13. The spoken input enables different possibilities of standardization and simplification; in this case a standard form of the cycle (*for*) command is used. This version is defined inside the system for the sake of simplification and it simplifies defining of the command (it supposes integer cycle variable and some other simplifications).

14. The system asks the user to comment the cycle. Such comments related to program blocks can be also used to provide position information.

24. The dialog system prevents the user to make errors in begin/end ({/} in C++) pair matching. The user must, however, signal to the system the end of actual program block when leaving it.

27-29. These dialog items demonstrate the behavior of the system when the user makes an error (erroneous "sm" instead of "sum")

35. The user asks for position information.

36. Position information includes the previously sampled comment (15).

The dialog generates the following C++ code:

```
integer sum, i;
sum = 0;
for(i=1; i<=1000; i++) {sum += i};   /* pointer to speech comment */
```

Even though the text form of the dialog is much longer then three simple lines of C++ code the example demonstrates some important advantages of the dialog system:

- Dominant speech communication;
- Intuitive dialog control;
- Simple access to variable declaration inside the source code;
- Suppressing of pair-matching errors and many other syntactical errors;
- Immediate access to position information;

Dialog generator of program source is presently developed for a fragment of C++ language for testing purposes. The further work on the generator will be coordinated with the Czech Union of Blind and Visually Impaired.

12 The Library of Procedures for Support of Programming of Visually Impaired People

The library consists of procedures and functions supporting work of visually impaired programmers;

- General routines supporting sound input and output;
- Procedures for speech synthesizer applications;
- Procedures for speech recognizer applications;
- General procedures for sound processing;
- Utilities for sound editing and mixing;
- Procedures and functions for semantical debugging support;
- Procedures for access and utilization of sound database;
- procedures for free form command input.

The library is supported by the speech synthesizer DEMOSTHENES ([2]) and speech recognizer RCG ([1]). Procedures for free form command input are supported by databases obtained by corpus analysis ([4]). The library is further supported by a sound database (samples of messages, earcons, environmental sounds, background sounds, signal etc.).

13 Application of the Blind User Oriented Hypertext System AUDIS

The hypertext system AUDIS ([3]) is primarily developed to provide comfortable accessibility to textbooks for visually impaired students. It pays maximal attention to the special needs of users and allows for comfortable customization. The system is fully supported by the utilities for speech communication and contains also utilities for text conversion and rendering. The system is developed and tested in close co-operation with the Czech Union of Blind and Visually Impaired.

We assume that AUDIS will be used also to provide overview of run-time libraries and to provide the information about the correct use of procedures and functions. The system should be also integrated into the dialog generator of program source code to support program building.

3. Programming in Logic Languages

There are several factors related to declarative paradigm that could be of some interest if we consider visually impaired programmers. We will demonstrate it on Prolog because it is typical, standardized, and widespread representative of logic programming.

Opposite to imperative languages, Prolog program is a composition of the unique type of data structures – terms. The meaning of a Prolog program is more intuitive as the components of the program are logical statements of a given form1.

The consequence for visually impaired users lies in the possibility to avoid treatment for various data structures and control structures, that makes the programming more natural and simple. However, the main difference issues from the style of thinking while programming. Imperative program is a sequence of steps that lead to the solution of a considered problem. In logic programming the computer is provided with author's knowledge about the problem (specification of the problem) in the form of facts and rules, thus enforcing insight of the problem and supporting high-level abstraction.

The knowledge should comprise a model of the problem. Well-designed model of the problem allows the user to ask different questions related to the problem, i.e. single program can give multiple solutions related to the problem.

Another benefits come from the inherited concepts of recursion and backtracking, simple manipulation with dynamic lists, powerful all-solution predicates, nondeterminism and many more. Newly added constraint solving features even allow to write programs just as a specification of the problem.

Dialog generator of the program clauses working on similar principle as in the previous section would be helpful for visually impaired programmers. We assume to analyze the relevant dialog and to make some experiments. Some examples comparing C and Prolog programs can be found at the following URL: *http://www.fi.muni.cz/usr/jergova/lp/comparisons.html*.

Object-Oriented Programming and Functional Languages

Although the basic ideas of object-oriented programming (modeling of the problem, encapsulation, inheritance etc.) seem to be very promising, this programming technology is effective especially for very complex program projects and needs high level support of integrated environment. The use of the object oriented programming can be hardly recommended without special integrated environment developed for blind programmers and if so, only for some special applications.

Since functional programming is based on the knowledge of Lambda calculus it is comprehensible only to limited audience of university students and some professionals. This limitation forces us to prefer imperative and logic languages.

Both object-oriented programming and functional programming are, according to our opinion, outside the sphere of real and effective possibilities for most of the visually impaired programmers.

5. Conclusions

In the paper we have discussed some problems that visually impaired encounter when programming. We have mentioned the project of the dialog generator of source code that is supported by the library of procedures for programming of visually impaired people and by the blind user oriented hypertext system AUDIS. The further work on this project will be directed both to implementation of new modules of the system and testing the system effectivity.

We have also briefly discussed the possible use of logic languages that seems to be (for selected applications) promising possibility for visually impaired programmers, since logic programming shifts the process of programming from the level of coding and thinking about the machine to the level of thinking about the actual problem (where the sight limitations may not play so significant role).

References

[1] KOPEČEK, I. (1997) Speech Synthesis of Czech Language in Time Domain and Applications for Visually Impaired. Proceedings of the 2nd SQEL Workshop on Multi-Lingual Information Retrieval Dialogues, Pilsen, pp. 141-145.

[2] KOPEČEK, I. (1997) Syllable Based Speech Synthesis. Proceedings of the 2nd International Workshop Speech and Computer - SPECOM'97, Cluj-Napoca, pp. 161-165.

[3] KOPEČEK, I. (1998) The Architecture of the Blind User Oriented Hypertext AUDIS. FI MU Report Series, FI MU RS-98-01.

[4] PALA, K., RYCHLÝ, P., SMRŽ, P. (1997) DESAM – Annotated Corpus for Czech. Proceedings of SOFSEM'97, Lecture Notes in Computer Science, Springer, pp. 523-530.

[5] DARVISHI, A. (1996) World Wide Web Access for Blind People: Problems, Available Solutions and an Approach for Using Environmental Sounds. Proceedings of ICCHP'96, Linz, pp. 369-373.

[6] STERLING, L., SHAPIRO, E., (1987) The Art of Prolog, Advanced Programming Techniques, MIT Press.

[7] STEVENS, R.D., WRIGHT, P.C., EDWARDS, A.D.N., BREWSTER, S.A. (1996) An Audio Glance at Syntactic Structure Based on Spoken Form. Proceedings of ICCHP'96, Linz, pp. 627-635.

An Assisting System For A GUI-Based PC For The Blind

Yutaka Shimizu[1], Masami Shinohara[2], Tadayuki Sasaki[3], Tetsuya Watanabe[4], Hiroki Minagawa[1] and Hideji Nagaoka[1]

Abstract

An assistive system for a GUI-based PC was designed for blind computer users. Both tactile and auditory devices were used for user interface. A tactile mouse and a tactile graphic display, designed for tactile use, and a speech synthesizer were introduced for auditory use. Utility software was developed for Windows 95, the most popular in Japan.

1. Introduction

With recent advances in the speech synthesizers and the Braille converter, it has become possible for blind persons to receive ordinary letter. However problems exist still, and further problems have appeared with advances in recent user interface.

One existing problem is the reception of a graphical image. Although the speech synthesizer is effective for reception of text, it is difficult to transmit non-linguistic information such as a graphs, charts and images. Tactile displays have been devised to overcome this difficulty. However there are few on the market.

Another problem is that of the GUI (graphical user interface). Although introduction of the GUI to the PC (personal computer) improved usability for sighted persons, it has confused blind users. A few screen-reading systems to access the GUI have began to be developed. In Japan they have not yet reached a complete solution.

The aim of the present study, is to design an assistive system to access the GUI type of PC and further to access a graphical image for the blind.

2. Design Approach

2.1 Present Condition for Blind Computer Users

Blind students at the Tsukuba College of Technology who selected the course for information processing usually use several PCs. Although assistive technology exists to access them, problems still exist. These are summarized in Table 1.

Although the Braille terminal and the speech synthesizer are available now, these are used to access text oriented applications. To access a graphic image tactile devices have been studied at several laboratories [1, 2, 3]. A challenging tactile display also has been developed recently [5]. Unfortunately, use of these are still limited at research laboratories because of cost and maintenance. An effort to access an OS (operating system) based on the GUI has yielded a

1) Tsukuba College of Technology, 4-3-15, Amakubo, Tsukuba, Ibaraki, 305-0005 Japan
2) National Institute for Bioscience and Human Informatics, 1-1, Higashi, Tsukuba, Ibaraki, 305-8566 Japan
3) Ibaraki University, 2-1-1, Bunkyo, Mito, Ibaraki, 310-8512 Japan
4) National Institute of Vocational Rehabilitation, 3-1-3, Wakaba, Mihamaku, Chiba, 261-0014 Japan

This research is supported by the Ministry of Education, Science, Sports, and Culture (grant-in-aid for scientific research #09558042).

screen-reading system named "the 95 Reader" in Japan [6]. Unfortunately, this utility can only be adapted to Windows 95 and matched to only a few applications.

Table 1.
Assisting state for blind computer user in Japan

object	state and problems
OS	DOS (possible), Windows and X-Window System(difficult)
word processor	Braille and speech synthesizer, but difficult for GUI-based application
spread sheet	speech synthesizer, but difficult to control each cell
database	Braille and speech synthesizer, but difficult to select items to be retrieved
drawing	impossible
DTP	impossible
E-mail	possible by the text mode, but difficult for GUI-based application
WWW	possible for text information, but impossible to access image
programming	possible for text-based language, but difficult for GUI-based application
computer education	difficult to teach by graphic image

2.2. Needs for Assistance

From the conditions mentioned above, the following seem to be needed for the present system.

- assistance of the GUI operation
- reception of graphic image
- simple drawing for graphics
- basic layout for the DTP
- reception of WWW information

2.3. User Interface

Recent PCs use the mouse frequently. There are two important functions for the mouse. One is a pointing and dragging function for the GUI-interface. Operation of the mouse by the blind has been examined to some extent [4]. The other is a drawing function to draw graphics. For sighted people, both functions work successfully.

Auditory and tactile modalities substitute for vision in this system. One tactile device, i.e., a tactile mouse that contains tactor-pins, is used as a control device. The tactile mouse presents partial information of the visual display. A second tactile device is a tactile graphic display. A spatial configuration of a visual screen is projected on the tactile display. In addition, synthesized speech announces linguistic information by a screen reader. Furthermore, the tactile display presents a graphic image.

3. System Design

3.1. Development Approach

Our goal of this study is not to develop a new PC for the blind but to develop an assistive system to the PC which can be used anywhere. The following have been adopted for the design of this system.

- Uses Windows 95, the most prevailing OS at present and uses Windows NT in near future.
- In conjunction with a Japanese screen reader, uses a tactile mouse and a tactile graphic display for operation of the GUI.
- Uses the speech synthesizer and/or the Braille terminal for reception of characterized information, and uses a tactile graphic display for not characterized information.

3i. Hardware Architecture

Figure 1 illustrates the hardware configuration is designed for this purpose.

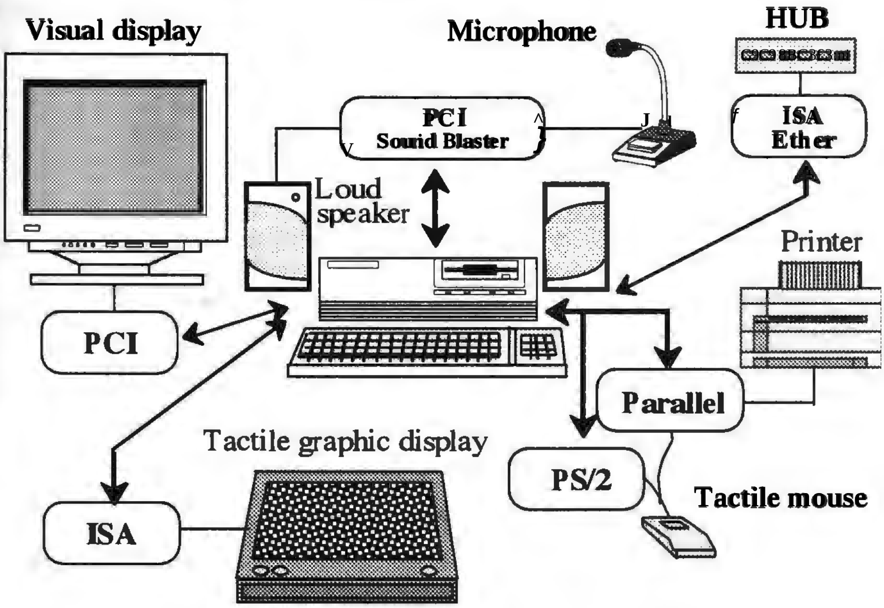


Figure 1. Hardware configuration used in this study.

Table 2.
Summary of bus-line for peripheral devises

Peripheral	Bus-Line
Visual	PCI (S3 TrioV 64+ standard)
CD-ROM	IDE (Mitsumi
Floppy Drive	Diskette controller
LAN	ISA (3COM 3C509)
Printer	LPT1 port (standard)
Sound Board	PCI (Sound Blaster AWE32)
Tactile	Mouse port (standard) and LPT1
Tactile Graphic	ISA (parallel
Braille	COM1 (standard)

* to be connected soon

The DELL's PC (Pentium Pro 180 MHz) with 64 MB of RAM is used. Synthesized speech app... from the sound board by the 95 Reader. The standard mouse port and the LPT1 port are used for the tactile mouse. An ISA bus is used for the tactile graphic display.
Bus lines used in this system are summarized in Table 2.

Tactile Mouse

The basic mechanism of the tactile mouse contains the PS/2 compatible mouse (MS InteliMou...) well as tactor-pins. An ergonomic effort has been pursued to fit the palm and the fingers. Activating mechanism for the tactor-pins is shown in Figure 2. The tactile surface consists of a 8 arrangement of tactor-pins with 3 mm interspacing. A bit-map projection of a visual screen presented on the pin array and a blind user recognizes it by the index and the middle fingers.

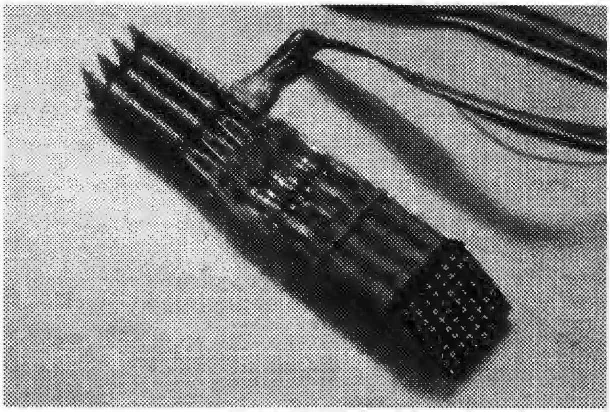


Figure 2. Picture of pin-array used in the tactile mouse.

Tactile Graphic Display

This device is the most expensive in the system. A new display will be developed next year with research funds. However, the previous 3-D display [5] is available now. In order to overcome the cost and the maintainability, specifications of a new display to be developed are simplified. These are compared in Table 3.

Table 3.
Tactile graphic display for this system

specification	3-D tactile display	new atactile display
diameter of tactor-pin	2.5 mm	2.5 mm
interspacing of tactor-pin	3 mm	3 mm
maximum height of tactor-pin	10 mm	2 mm
number of elevating levels	100 levels	0 or 1
number of picture cells	64 x 64	excess 128 x 64
presenting time	5 s/frame	2 s / frame

3. Software Architecture

Figure 3 illustrates the software architecture defined for the hardware configuration. The 95 Reader is the auditory display, whereas the tactile mouse is the tactile display for the GUI access. Utility software has been developed for the tactile mouse. This software feeds bit-map information of the visual display to the tactor-pins. Gray scale of density information on the screen converts to 0 or 1 at any adjustable threshold for the tactor-pins. Presenting area of the visual screen is also converted into a minimum of 8 x 8 pixels. For reception of WWW, a commercially available software (GANNOSUKE) that presents Japanese voice reading is used. The Via Voice (IBM) is also used for Japanese speech recognition. Utility software for the tactile graphic display is now being developed.

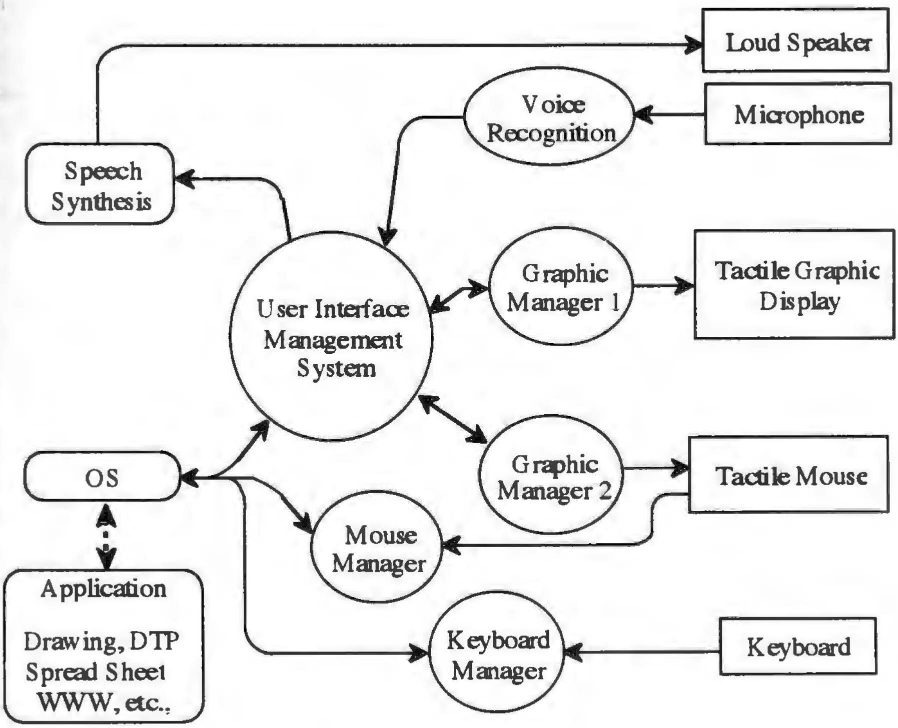


Figure 3. Software architecture.

4. Concluding Remarks

A PC system by integrative use of sound and tactile devices was designed. This system seems to offer accessibility for the GUI operation and for the reception of graphical images. Further study is warranted to improve the system and for evaluation of acceptance.

5. References

[1] FRIKSEN-GIBSON, S.F., BACH-y-RITA, P., TOMPKINS, W.J., and WEBSTER, G.J., A 64-solenoid four-levdi fingertip search display for the blind. IEEE Trans. Biom. Eng., BME34, pp.963-965, 1987.

[2] SCHWEIKHARDT, W., and KLOPER, K., Rechnerunterstutzte aufbereitung von eirdshirmtxt-grafiken ineine tastbare darstellung, in; Institut fur Informatik Universitat Stuttgart, pp.1-16, 1985.

[3] SHIMIZU, Y., Tactile display terminal for visually handicapped. Displays, July, pp.116-120, 1986.

[4] SHIMIZU, Y., SHINOHARA, M., and NAGAOKA, H., A tactile display for mouse operation by blind compota^ user, in; J.Klaus, E.Auff, W.Kremser, and W.Zagler (Eds.) Interdisciplinary Aspects on Computers Helping People wili Special Needs, pp.755-760, 1996.

[5] SHINOHARA, M., SHIMIZU, Y., and MOCHIZUKI, A., Development of a 3-dimensional tactile display for the blind, in; Proceedings of the ECART 3, pp.404-405, 1995.

[6] WATANABE, T., OKADA, S., How to minimize the selecting time from auditory menues, in; J.Klaus, E.Auff, W.Kremser, and W.Zagler (Eds.) Interdisciplinary Aspects on Computers Helping People with Special Needs, p.791, 1996.

GONVI

Non-Visual Access to Documents and GUIs with a Constraint-Based Approach

Gottfried Zimmermann[1]

University of Stuttgart, Germany

mRACT

JMspaper describes a constraint-based interactive system for non-visual representation of graphi- ddocuments and graphical user interfaces (GUIs). GONVI („Graphical Object Server for Non- \lal Interaction“) establishes spatial and other constraints between textual and graphical ob- fcts; these constraints rule the transformation process for non-visual interaction. Of central inter- tsisthe integration of text and graphical elements considering the non-scalability of braille. A prototype of GONVI is currently being implemented in Java using PostScript documents as graphi- cal information source.

1 troduction

Graphical information plays an important role in communications. Most prints like school books, office documents or instruction manuals today contain images and sketches which are essential for understanding. Today's human-computer interaction bases on graphical interfaces and visual inter- [?] techniques mainly designed for sighted users. One of the main electronic information sources [?] society, the World Wide Web (WWW), also exhaustively makes use of graphical elements which are mostly not sufficiently described in the textual context and therefore cannot be regarded redundant.

In general graphical images and sketches are interwoven with textual elements that form an integral part of the graphical representation. The graphical and textual elements with their mutual spatial relations raise a main problem for blind readers and computer users. Graphics can be transformed to tactile images to a certain degree, and text is traditionally converted to braille for non-sighted readers. But the fixed size of braille characters in contrast to the scalability of ink-print fonts causes either huge tactile formats or at least the loss of spatial relations implied by the visual original.

[1] Email: gottfried.zimmermann@t-online.de

There are several approaches to solve this problem. [Lötzsch94] mentions four types of tactile graphics: 1. full text graphics possibly with pointers, 2. labelled graphics (with additional legend) possibly with pointers, 3. textual guided pure graphics (with an additional guide in braille or on tape) and 4. pure graphics (requires help by a person or by a computer). The production of tactile graphics still remains a time-consuming process requiring manual design by experts for each graphical image.

In the non-visual representation of graphical user interfaces the so called screen reader programs generally concentrate on textual elements by speech output or output on a braille terminal, but more or less neglect the reproduction of graphical elements and their relations to the interwoven text parts [Burger/Stöger96]. This leads to a non-acceptable loss of information for blind computer users.

For the non-visual representation of text and graphics we need a solution that provides

- non-visual access to (almost) every information given in the visual original,
- extensive preservation of spatial relations between textual and graphical elements,
- the possibility of controlling the transformation process with respect to domain-specific requirements and user skills,
- an automatic and reproducible transformation process in general, but the possibility of manual intervention in certain cases,
- different results for different user environments,
- type-specific (object-oriented) interaction techniques for graphical and textual elements,
- the availability of different braille modes (6-dot-braille, 8-dot-braille, contracted braille), and
- applicability to printed/electronic graphical documents *and* graphical user interfaces (GUIs).

2. Constraints

2.1. Constraint Programming

A *constraint* is a declarative description of mutual relations between a set of objects. Constraints affect the properties of objects on runtime in correlation to other objects' properties. In contrast to the assignment in imperative programming languages which only defines a one-sided relation at evaluation time, a constraint affects all involved objects from creation until deletion time of the constraint. Constraints do not specify an algorithm for the determination of the constrained properties – it is the job of a *constraint solver* to satisfy as many constraints as possible. The constraint solver enforces a constraint every time an involved object has changed.

A set of constraints may include conflicting constraints which cannot be satisfied all together. To specify an order in which the constraints should be solved constraints can be organised as *constraint* hierarchy. The most important constraints belong to the highest hierarchy level. The constraint solver guarantees that when solving a constraint of a certain hierarchy level no other constraint of a higher level is violated.

2.2 Constraints and Their Use in GONVI

In GONVI constraints represent mutual relations between graphical objects[2] or relations between graphical objects and an output or input device. Constraints set restrictions to the representation and interaction process in order to access the graphical objects. This means that the system's behaviour is controlled by a set of graphical objects, a set of output and input devices and their associated constraints. Since in GONVI constraints are created by user-defined rules the interaction of the system has to be specified by implementing a set of constraint creation rules.

2.2.1. Spatial constraints

To maintain spatial relations in a visual graphical document or interface these relations have to be detected and expressed by *spatial constraints* between graphical objects. The spatial constraints include:

- common horizontal or vertical justification of graphical objects,
- minimal or maximal distance, e.g. between a text and an arrow tip, and
- spatial inclusion, e.g. a text in a frame.

2.2.2. Medium-specific constraints

Medium-specific constraints specify the non-visual representation in the context of a given user environment. These constraints are generally given by technical restrictions of the output or input medium. Medium-specific constraints define:

- The size of braille characters. This depends on the resolution of the tactile output medium. E.g. braille printers have a finer dot resolution and therefore allow a greater „braille step length" than dynamic braille displays.
- Filtering of object properties. E.g. the property colour cannot directly be displayed on a tactile output device.
- Restrictions of information density on an output medium. In dependence of resolution and scale factor of a given output medium fine-granular information has to be neglected in favour of the whole context.
- Filtering of graphical objects. Not every graphical object is suitable to be displayed on a given user environment.

---

[2] In this paper the term „graphical object" refers to both graphical and textual objects.

2.2.3. Interaction constraints

Interaction constraints directly affect the way the system interacts with the user. They specify „interaction rules" like

- braille mode for text objects (6-dot-braille, 8-dot-braille, contracted braille, labels), or
- controlling the output focus, e.g. selecting some objects to be zoomed in by dragging a pointing device.

3. GONVI System Architecture

GONVI is a constraint-based transformation and interaction process. It takes a set of graphical objects (from a visual context) as input and provides representation and interaction techniques for a non-visual user environment as output. In the GONVI prototype (see section 4) this process is implemented for a typical graphical information source (input) and user environment (output) (see *Figure 1*).

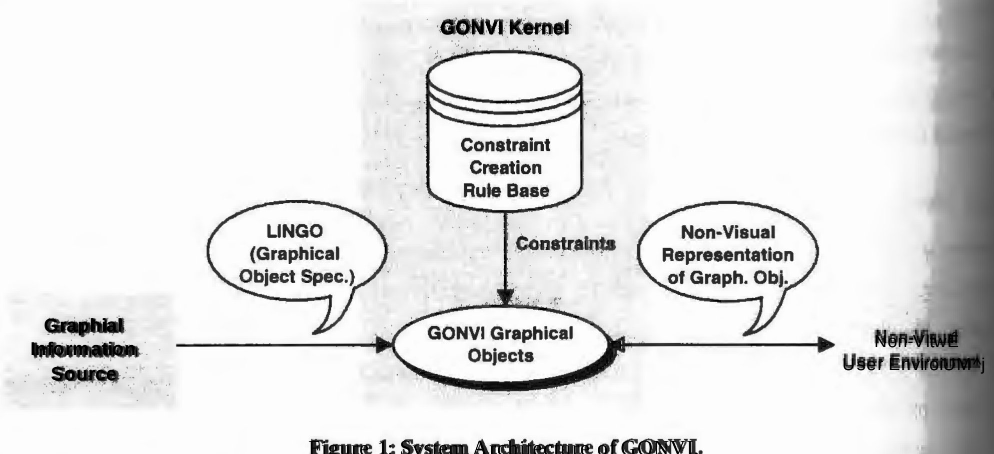


Figure 1: System Architecture of GONVI.

3.1. Lingo

GONVI defines a special language for graphical object descriptions. Each graphical information source encodes its graphical description in *Lingo* („*L*anguage for *In*formation Exchange on Graphical Objects"). Since Lingo is used in the context of inter-process communication (IPC) and communication through character streams, it is a character-based language.

A special Lingo dialect is introduced for each graphical information source. On the other side, the GONVI kernel provides a special parser component for each Lingo dialect. In GONVI, builder objects [Gamma et al. 95] are responsible for building internal graphical object instances from lUlgo dialect descriptions.

| u! GONVI Graphical Objects

(ONVI istinguishes three types of graphical objects. *Graphical primitive objects* are the atomic ... of graphical or textual information; examples are straight line segments, curve segments, ...cters, filled areas or pixel-based bitmaps. *Graphical composite objects* consist of graphical ...itive or other graphical composite objects and represent complex graphical descriptions which ... be regarded as a greater unit; examples are rectangles, tables or text paragraphs. For application-...cific contexts GONVI provides *graphical application objects* which consist of graphical primi-... or graphical composite objects. Graphical application objects represent semantic units of an ...plication's domain or the real world; examples are interaction objects (e.g. buttons, menus) or real ...ld objects.

...sing a Lingo description results in a set of (mostly) graphical primitive objects. In GONVI a *...phical object detector (see Figure 2)* creates higher level graphical objects (graphical composite ... and graphical application objects) on top of graphical primitive objects.

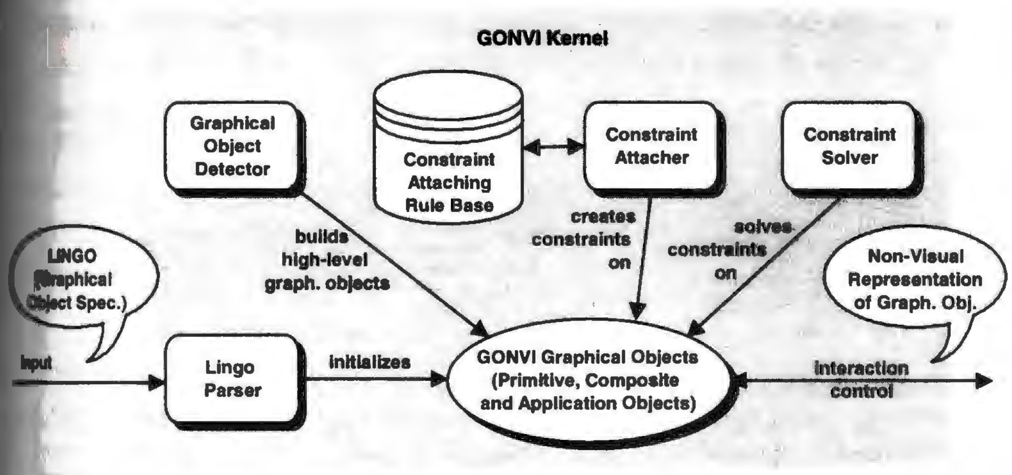


Figure 2: Internal components of the GONVI kernel.

3. Attaching Constraints by Rules

...constraint attacher (see Figure 2) seeks conditions for creating and attaching constraints (see sec-...ion 2) between the graphical objects. Medium-specific constraints also consider the complete user environment with its output and input devices. This process is triggered by the *constraint attaching rule base* which stores user-defined rule sets. They can be regarded as domain and environment-...cific interaction descriptions. For experimental purposes the user can inspect the automatically

created constraints and additionally attach constraints to individual graphical objects and/or environmental devices. The result of this phase is a set of constraints organised as constraint hierarchy.

3.4. Constraint Solving

After creating the constraints a *constraint solver (see Figure 2)* has to satisfy them. Many different constraint solving techniques are described in literature, e.g. in [Leler88] and [Saraswat/Van Hentenryck95]. In GONVI the constraint solver must be able to handle constraint *hierarchies*. Because of its applicability to GUIs it also must be *incremental*.

3.5. Non-Visual Representation of Graphical Objects

In the object-oriented design of GONVI each type of graphical objects can implement its specific representation and interaction techniques. Thus tables can display a cell's content on a dynamic tactile device while on the same time the corresponding row and column heading are given by speech output, for instance.

4. Prototypical Implementation of GONVI

In our prototypical implementation the description of graphical objects is taken from a description in PostScript format [Adobe88]. For non-visual interaction the following devices can be connected: tactile output devices (pin matrix display and/or tactile printer), speech and sound output.

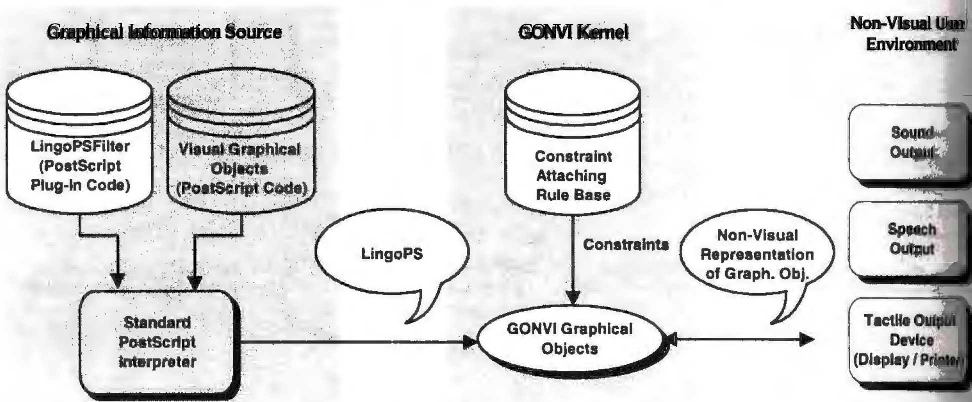


Figure 3: System architecture of the GONVI prototype. The white components are direct parts of the prototypical implementation.

4.1. Lingo PostScript Filter

For the GONVI prototype a PostScript program called *LingoPSFilter* implements a „plug in" for any PostScript interpreter in order to generate a graphical description in the *LingoPS* format from a given PostScript description. LingoPSFilter must have been run in the PostScript interpreter prior to the actual PostScript description. It redefines some of the PostScript system operators and thus reproduces the „whole picture" in the LingoPS language, a special Lingo dialect used for PostScript information sources.

[...kenhäuser/Weber94] introduce three main approaches for accessing graphics and graphical user interfaces for blind users: the bottom-up, middle-out and top-down approach. The GONVI prototype with its PostScript information source implements a mixture of the bottom-up and the middle-out approach. The main reasons for using PostScript as graphical information source are:

- PostScript is a vector-based, device-independent description language for text and graphics. In PostScript texts are coded as strings, so there is no possible loss of information through optical character recognition (OCR).
- Since 1982 PostScript became an industrial standard for graphical printer devices. Moreover DisplayPostScript, an extension of PostScript, is used for controlling the graphical user interface in some unix-related operating systems (e.g. Nextstep, Openstep). This indicates that it is possible to have a common graphical description language base for both displays and printers.
- Many existing documents are coded in PostScript. They are presently (almost) inaccessible for blind persons. In context of hypertext systems like online help manuals there is also a growing interest in documents coded in the PostScript-based Portable Document Format (PDF). Principally every paper document can be converted into the PostScript format by OCR and vectoring software.

The present prototype of GONVI is a very special application whereas the GONVI kernel is basically independent from any concrete graphical information source. So it is possible to implement the GONVI system to use graphical information delivered by the Hypertext Markup Language (HTML) or the Java Abstract Window Toolkit (AWT), for instance.

4.2 The User Environment for the Prototype

The present user environment focus lies on *tactile output;* in this medium spatial relations can be most adequately expressed. Tactile output can be given on a tactile printer or on a dynamic pin matrix with an integrated pointing device, the latter allowing real user interaction techniques like zooming in/out and user-driven exploration of graphical properties.

Speech and sound output can be used as additional devices to support different interaction modes. The use of speech and sound is driven by the constraints attached to the graphical objects and therefore depends on the constraint attaching rule base (see section 3.3).

5. Conclusion

In our world in which information more and more consists of graphical representations designed for visual perception there is a great need to make this information also accessible to non-sighted persons. GONVI describes and shows how graphical information can be transformed into non-visual, representations without loosing its internal consistency. The powerful concept of constraint programming also allows specifying interaction techniques for a non-visual user environment.

6. References

[Adobe88] Adobe Systems Inc.: PostScript – Handbuch. Addison Wesley, Bonn, 1988.

[Burger/Stöger96] F. Burger, B. Stöger: Access to the WWW for Visually Impaired People using Browsing Tools based on GUI: State of the Art and Prospects. In: ICCHP'96 Proceedings, Oldenbourg, 1996, pp. 303–310.

[Gamma et al. 95] E. Gamma, R. Helm, R. Johnson, J. Vlissides: Design Patterns – Elements of Reusable Object-Oriented Software. Addison-Wesley, 1995.

[Gunzenhäuser/Weber94] R. Gunzenhäuser, G. Weber: Graphical User Interfaces for Blind People. In: K. Brunnstein, E. Raubold (Eds.): 13th World Computer Congress 94, Vol. 2. Elsevier Science B. V., Holland, 1994, pp. 450-457.

[Laufenberg/Lötzsch95] W. Laufenberg, J. Lötzsch: Tastbare Abbildungen für Blinde – Thesen zu Bedarf, Entwurf, Fertigung, Zugriff. In: W. Laufenberg, J. Lötzsch (Eds.): Taktile Medien – Kolloquium über tastbare Abbildungen für Blinde, Tagungsband, Deutsche Blindenstudienanstalt e.V., Blinden- und Sehbehinderten-Verband Sachsen e.V., Nov. 1995, pp. 11–22.

[Leler88] W. Leler: Constraint Programming Languages – Their Specification and Generation. Addison-Wesley, 1988

[Lötzsch94] J. Lötzsch: Computer-aided Access to Tactile Graphics for the Blind. In: ICCHP'94 Proceedings, Springer, 1994, pp. 575–581.

[Saraswat/Van Hentenryck95] V. Saraswat, P. V. Hentenryck (Hrsg.): Principles and Practice of Constraint Programming. MIT Press, 1995.

[Wisskirchen90] P. Wisskirchen: Object-Oriented Graphics – From GKS and PHIGS to Object-Oriented Systems. Springer, 1990.

7. Acknowledgement

I wish to thank Prof. Dr. Gunzenhäuser and many members of his research group for help and inspiring discussions.

OPERATIONAL HELPING FUNCTION OF THE GUI FOR THE VISUALLY DISABLED USING A VIRTUAL SOUND SCREEN

Michio SHIMIZU[1)], Kazunori ITOH[2)] and Yoshimichi YONEZAWA[2)]

Abstract

We already proposed a handwriting support system for the visually disabled by using a virtual sound screen. We confirmed that the support system is helpful when adventitious blind people write characters and draw figures. Then, the drawing support system was developed in the MS-windows environment of a personal computer. Some existing problems, such as leading a mouse or a pen cursor, still remain. In this paper we examine the possibility of operational support in the GUI environment for the blind using a virtual sound screen. As the result of many considerations, we found that a virtual sound screen could be applied to leading a mouse cursor, in addition to the drawing support.

1. Introduction

The population of the visually impaired in Japan is about 330,000 and this rate rises each year due to the addition of the aged. The number of the early blind has decreased due to the progress of medical techniques, but adventitious blindness, resulting from traffic accidents and diabetes has a tendency to increase. The aged acquired blind would like to continue writing letters and drawing pictures as sighted people do.

Recently, computer user interfaces have been realized using graphical presentation that imitates a real desktop environment so as to simplify user operation by clicking on graphical objects with a pointing device. These kinds of interfaces are convenient for sighted people, but for the visually disabled it has become more difficult to use computers over the past few years. In earlier systems the user interface was a simple character based environment. In those systems, blind computer users could access the command-line based interface by using a screen reader which usually read the text displayed on the screen and provided information to either a speech synthesizer or a tactile text display.

1) Department of Liberal Arts, Nagano Prefectural College, 8-49-7 Miwa Nagano-shi, Japan. E-mail: ito@ns.nagano-kentan.ac.jp. 2) Department of Information Engineering, Faculty of Engineering, Shinshu University. 500 Wakasato Nagano-shi, Japan. E-mail: itoh@cs.shinshu-u.ac.jp, yanezawa@cs.shinshu-u.ac.jp

In order to help the acquired blind with handwriting, we already proposed the sound support system which converted a pen position into a point of sound image by using control of sound localization [7,8], Handwriting patterns can be expressed in a virtual sound screen reproduced on headphones Recently, several non-visual interface system have been proposed [1,6,10] and a GUI-based screen reader has been developed by a spatial audio processing system with synthesized speech output [2,5]

We also previously proposed the drawing support system for the visually disabled by using a virtual sound screen [9], The same idea might be applied to the support of operation in a GUI circumstance for the visually impaired computer users. This paper deals with the construction of a virtual sound screen based on the experimental factors of sound location. Position of a mouse cursor is converted into a point sound image, which is synthesized by the factors of sound localization.

We examine the possibility of the operation support in the GUI environment for the blind using a virtual sound screen. The purpose of this paper is confirming the reliability of leading a mouse cursor to any position in the window by use of a virtual sound screen. As the result of many considerations, we found that the virtual sound screen could be applied to leading a mouse cursor in addition to drawing support.

2. A virtual sound screen and it's application to drawing

When a sound is coming from some distance away and from a particular angle to the listener, there are a number of cues that can serve to indicate the direction of the sound source. We determine the locations of sound sources with differences in auditory information received by listening with two ears. The location factors of the horizontal plane are interaural level differences and interaural time differences between both ears [3]. It is not clear what elements contribute to the localization for the median plane. Sound reflection and scattering on the pinna are important cues of location.

Moreover, low frequency sounds tend to be perceived the image to a low position, whereas for high frequency sounds the situation is reversed [4]. The computational simulation of spatial sound images was investigated by the head related transfer function (HRTF) [11] or the pinna related transfer function (PRTF). Therefore, the simulation of sound localization by listening on headphones can express the point where sound images correspond to the picture elements.

Figure 1 shows a virtual sound screen which is composed of 24×24 point sound images synthesized by factors of sound location. The horizontal axis of the screen is composed by interaural level differences between right and left ears. Maximum difference is set at 14 dB. The vertical axis of the screen is composed by a psychological factor due to frequency alteration. Each display point is based on the musical scale whose lowest frequency is 400 Hz and highest is 2000 Hz. We have

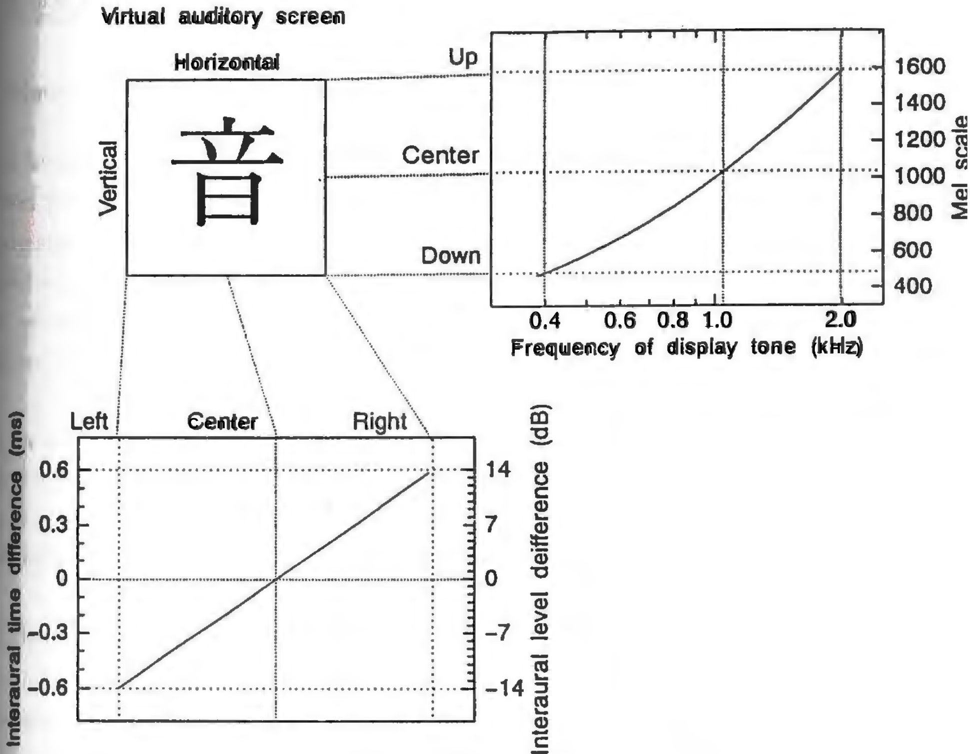


Fig.1 A virtual sound screen of the support system.

nstructed a vertical axis of the sound screen with the pinna related transfer function.

Figure 2 shows a block diagram of the support system for handwriting and drawing by a mouse or a tablet. The resolution of the CRT window display is 640×480 pixels and the client area of the window corresponds to a virtual sound screen. When we move a mouse cursor or pen cursor, two dimensional data of positions are transferred from the pointing device to the personal computer from the sound board (Creative Media Sound Blaster 16) and are converted into brief tones of 30 ms display duration and about 10 ms silent duration. Finally, these signals are amplified and sent to headphones.

The developed program acts as event driven. When an event such as the movement of a mouse cursor or timer interruption occurs, the windows system sends a message to the program and the program obeys it. Clicking a button of the mouse in the client area of a window, the program then enters painting mode and the mouse cursor is captured only within the client area, as the movement of the mouse and the positions of the mouse cursor are taken into the personal computer in the fixed time interval. The point sound images and the graphical images corresponding to the positions of the mouse cursor are displayed. Data of image is stored as a file, with the ability to be recalled to display at any time.

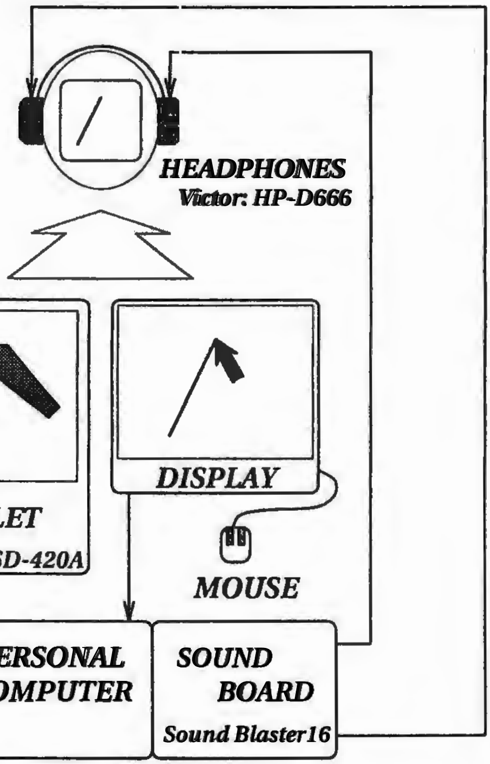
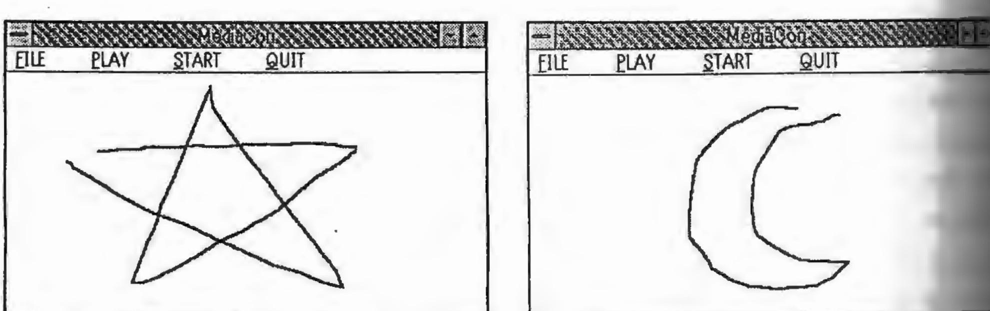


Fig.2 Block diagram of the support system.



(a) star

(b) moon

Fig.3 Examples of outline drawings.

We examined what happened when the acquired blind drew arbitrary patterns utilizing a mouse. Figure 3 shows the examples of outline drawings by one subject using this system. The drawn figures are well balanced on the windows.

3 Mouse cursor-leading method

It was found that the drawing support system developed on the windows is helpful for the visually impaired subject experimentation [9]. However, moving the mouse cursor to the start point in the drawing area or menu operation such as the recording and replaying of the drawn information are also required in addition to the drawing operation in order to completely utilize this application software. Then, we have to consider the operational support function that leads a mouse cursor to any position on the screen to be operated only by the blind user.

Now, we suspect that it also may be possible to lead a mouse cursor by using a virtual sound screen because the visual screen of personal computer corresponds to a virtual sound screen; therefore, we tried to examine the possibility through our a basic experiment using a model that is made from twenty buttons on a screen.

An explanation of experimentation criteria:

Screen lay-out: Twenty buttons with a size of 7 X 7 pixels are arranged on a 400 X 640 pixel screen as illustrated in Figure 4.

Subjects: Five university students, A, B, C, D and E, possessing normal vision. Each subject is blindfolded for the duration of the experiment. Two subjects, A and B are trained in mouse operation while C, D and E are novices.

Sound display: Brief tones of 30 ms duration and 10 ms tone pauses are alternately repeated. It has been shown that this display method is best to effect sound localization [8].

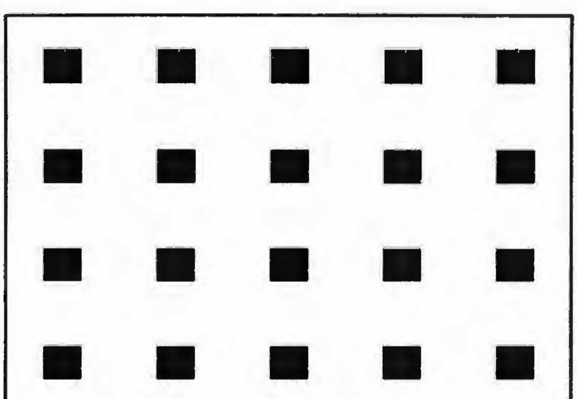
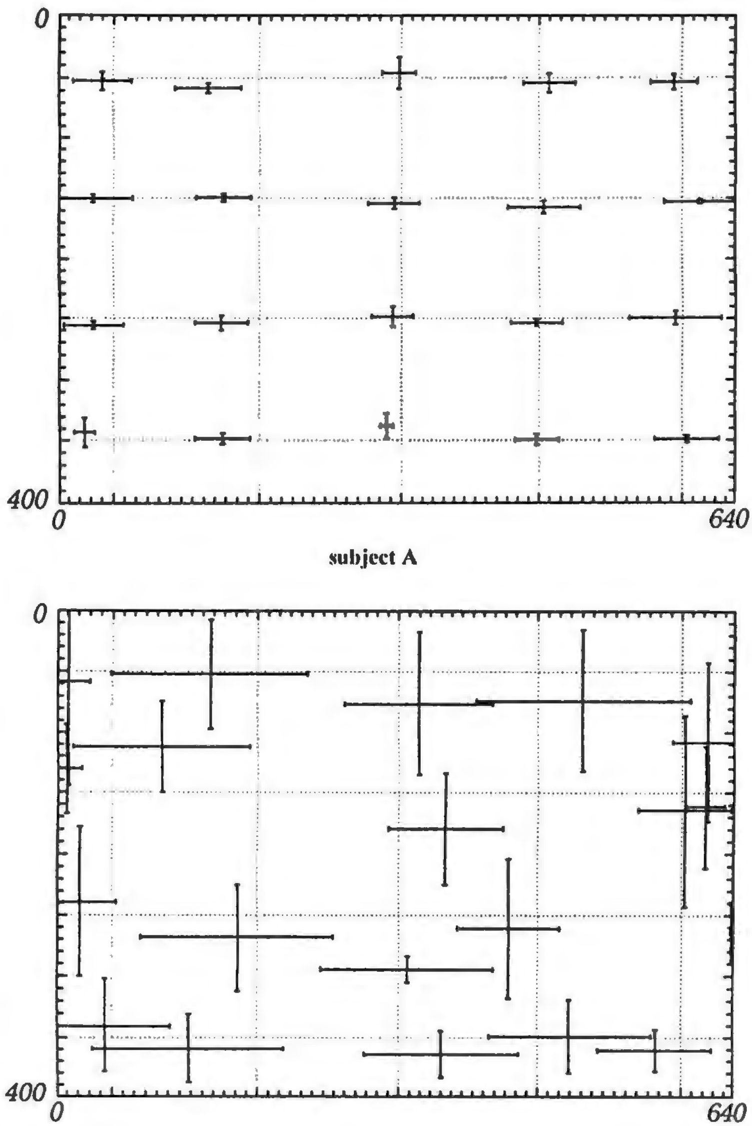


Fig.4 Screen lay-out.

A preliminary experiment was simple. The subject tried to move a mouse cursor to the sound display point on the selected button through using the virtual sound screen. By the method described below, each subject was tested 100 times,

STEP 1: Randomly choosing one button from twenty buttons, the corresponding sound of the position are displayed.

STEP 2: Subject tries to move a mouse cursor to the button by hearing the sound that corresponds to the position of the mouse cursor.



subject C

Fig.5 Result of experiment.

STEP 3: When subject decides that the mouse cursor might be on the designated button, he clicks the left button of the mouse immediately.

Figure 5 shows the result of the experiment for subject A and C. In this figure, the mean value and the standard deviation of the click point for each button are presented. The difference between the click point and the actual display point is so large that it may be difficult to move a mouse cursor to the button in such a simple manner.

After several investigations, it was found that the following method may be the most effective to lead a mouse cursor.

STEP 1: Randomly choosing one button from twenty buttons and the corresponding sound of this position is displayed.

STEP 2: Subject tries to move the mouse cursor to the button by hearing the corresponding sound of the position of the mouse cursor. Whenever he would like to confirm the position of the button as object, he can hear the sound for the button by clicking the right button of the mouse. If the mouse cursor is out of the window, white noise is produced.

STEP 3: When the mouse cursor is moved onto the designated button, the sound changes. As soon as the subject confirms it, he clicks the left button of the mouse immediately. The changed sound is given as an overtone of the original sound and as well as cutting the volume in half.

The experimentation procedure was conducted 100 times, with 20 buttons per window, each chosen at random 5 times. We evaluated the mean value and the maximum value in the leading time, that is the time from the end of STEP 1 to the click in STEP 3.

Table 1. Leading time.

subject	mean(sec)	maximum(sec)
A	8.67	22.03
B	4.81	15.18
C	24.53	117.06
D	21.06	117.35
E	23.75	123.56

The results of the experiment for five subjects are shown in Table 1. As for subjects A and B, who had already mastered mouse operation, it was possible to move the mouse cursor within 10 seconds on the average. However, since subjects C, D and E were novices movement within 30 seconds the average, but took two minutes in the worst case. We observed that the leading time became sh in the process of 100 attempts. Then, it may be concluded through the results of this experime leading time and skill of operation are fairly related.

4. Conclusion

In this paper, we examined whether or not it's possible to apply a virtual sound screen to the sup of operation in the GUI circumstances for the visually impaired. As a result, in addition to drawing support, it was proved that the virtual sound screen could be applied to the mouse ar leading method as one of the operational support. For future study, we would like to conduct seve experiments upon subjects who are visually impaired or by using various sounds or screen lay Up to now, GUI for the blind was assumed to be a stand-alone system, but it may be able to be to usability on the Internet Web.

Acknowledgement

The authors would like to thank Mr. Norio Kitamura, Mr. Masami Hashimoto, Mr. Yoshihiro Ina and Mr. Hiromasa Kaneko for their helpful discussions. This work was partly supported by a Gr in-Aid for General Scientific Research from the Japanese Ministry of Education, Science and Cul

References

[1] Agranovski A.V., Evreinov G. E. and Yaslikin A.S.: Graphic audio display for the blind. Proceeding of ICCHP (1996) 539.

[2] Aritsuka T. and Hataoka N.: GUI Representation System Using Spatial Sound for Visually Disabled. Proc of ASWA97 (1997) 415.

[3] Blauert J.: Spatial Hearing: The psychophysics of Human Sound Localization. MIT Press. Cambridge, MA

[4] Blauert J.: Sound localizing in the madian plane, Acustica, 22 (1969/70) 205.

[5] Crispien K., Wurz W.and Weber G.: Using Spatial Audio for the Enhanced Presentation of Synthesized Spe within Screen-Readers for the Blind Computer Users, Proceedings of 4th ICCHP (1994) 144.

[6] Edwards A.D.N.: Sound track: An Auditory interface for the Blind Users. Human-Computer Interaction (1989) 45.

[7] Itoh K. and Yonezawa Y.: Support system for handwriting characters and drawing figures to the blind feedback of sound imaging signals. J. Microcomputer Applications. 13. 2 (1990) 177.

[8] Itoh K., Inagaki Y. Yonezawa Y. and Hashimoto M.: A support system for handwriting for the blind using a sound screen, Symbiosis of Human and Artifact, Elsevier, Amsterdam (1995) 803.

[9] Itoh K., Michio S. and Yonezawa Y.: A drawing support system for the visually impaired computer user virtual sound screen, Design of Computing Systems. Elsevier. Amsterdam (1997) 401.

[10] Kirschner A. I.: Navigating the Graphical User Interface (GUI) by the Visually Impaired Computer Symbiosis of Human and Artifact, Elsevier, Amsterdam (1995) 149.

[11] Wightman F.L. and Kistler D. J.: Headphone Simulation of Free-Field Listening II: Stimulus Synthesis. J. Soc. Am., 85 (1989) 858.

Human-Computer Interaction (2)

GENERATING ON-LINE CROSS-REFERENCED HCI GUIDELINES FOR THE DISABLED

Luc Goffinet, Monique Noirhomme-Fraiture [1]

Abstract

This paper deals with the problem of converting existing human-computer interface guidelines for the disabled into hypertext on the WWW. First, it introduces a method to generate hypertext from text. This method relies on the structuring of an ASCII text file to generate hierarchical and navigational links. It also relies on statistical measures to generate cross-references. Then it goes on with a case study concerning the hypertext conversion of general HCI Guidelines. Finally it tackles the problem of specific HCI disability guidelines, by showing how four different sets of guidelines have been integrated into a single on-line hypertext database.

1 Introduction

Guidelines for human-computer interaction are abundant in scientific literature. Famous examples include [15] and more recently [10], HCI Guidelines also exist for people with disabilities, though more sparingly. Here are some of them : [17], [9], [8], [18], [19]. European research projects such as INCLUDE or HUSAT aim at bringing together the pieces of the jigsaw puzzle.

Up to now, most of these guidelines are available in electronic form, and some of them even exist in a hypertext format (see [21]). But, to our knowledge, very little concerning HCI and the disabled has been turned into full-blown hypertext yet. Except maybe the INCLUDE project about Public Access Terminals [5].

Having on-line cross-referenced HCI guidelines would of course be helpful to software designers, as guidelines dealing with the same aspects of HCI for the disabled are currently scattered among different sources, making it hard for people to find relevant advice in the field [16]. The hypertext cross-references within different documents on the WWW covering the same area would be a major improvement too. These are the issues that we shall try to address in this paper.

2 From text to Hypertext

The problem of converting text into hypertext is mainly twofold. The first part of it is the generation of *hierarchical links*, such as "table of contents" references, "next", "previous", "top of chapter"

[1] Institut d'Informatique, 21 rue Grandgagnage, B-5000 Namur, Belgium.
Phone : +32/81/72.49.93. Fax : +32/81/72.49.67. E-mail: {lgo,mno}@info.fundp.ac.be
http://www.info.fundp.ac.be/~lgo/

buttons, etc. This is of course the easiest part of the translation process. For example, generating several HTML documents from a single well-structured Word document only requires a good set of macros. More generally the problem of turning regularly structured text into hypertext has already been addressed by many researchers (see [3, 12, 14]).

The second part of the conversion process is the automatic generation of cross-reference links without any textual clues (such as « see also » or « cf »). So far this has largely been done by hand, by people that have sufficient knowledge about the domain and who can point out *semantic links* between different parts of a document. Automatic cross-referencing is a challenging task, because hypertexts tend to grow in size and this task is highly time-consuming.

After a brief introduction to the methods that we have used to generate hypertext (thanks to a real world case study), we shall tackle the specific problem of turning HCI guidelines into hypertext.

2.1. Generation of hierarchical links

The most obvious part of the hypertext conversion process is thus the generation of a bunch of HTML files, all related to one another according to their spatial order (first, next, previous, last, top, bottom). All that is required as input to this process is a well-structured source text, where the hierarchical structure is marked by styles attached to paragraphs (such as "Heading 1", "Heading 2", etc... in a word processor).

We have used the RTFtoHTML software [11] to translate the text from the well-known RTF format into a hierarchy of well-structured HTML files. This also implies that the output of the process has always the same structure, which ensures consistency across different guideline sets.

2.2. Generation of semantic links

The remaining task has been to generate meaningful cross-reference links that can help people gather useful knowledge about related items without browsing through the entire collection of guidelines.

There are in fact several approaches to detect cross-reference links. The first one is the *syntactic approach*. Cross-references can be detected by occurrences of phrases such as : *"see also ...", "this can be liked to ...", "in ... Mister X says that"* or by more sophisticated mechanisms to detect quotations [1]. This approach is a first step towards the final goal. In the case of Smith & Mosier's guide, it has been followed by [21] to automatically generate an hypertext for Windows. This approach can be extended by *natural language analysis*, used by some researchers to detect hypertext 'anchors' [7].

The second approach comes from the Artificial Intelligence field. It is based on *neural networks*. As the description of this approach is beyond the scope of this article, the interested reader is referred to [6, 20].

The third approach is based on *statistical methods* used in Information Retrieval. When you search for a document in a database, a keyword vector is generated from your query and compared to the keyword vectors representing the documents inside the database. In a nutshell, a keyword vector consists of any significant word that belongs to a document. A similarity coefficient is then computed for every pair query/document. The best matches, i.e. those documents that have the higher similarity coefficient with the query, are given as output of the process. More details can be found in [13].

This is the basic idea behind the generation of cross-reference links based on the statistical analysis of a document. Every document in the database is compared with all the others and, when similarity coefficients are above a given threshold, a link can be generated.

3 A case study

We have picked up the guidelines from Smith & Mosier [15] to carry out our first experiment. A complete report of this study can be found in [4].

Their guide contains 957 rules for designing more usable software interfaces. Our goal has thus been to find a way to create meaningful *semantic links* between those rules. This work has also been made easier by the fact that Smith & Mosier had already inserted 701 such links in their guide. This has been important not only for evaluation purposes, but also to serve as a basis for optimised keyword weights used in probabilistic information retrieval.

We have chosen the statistical approach described above, using the *weighted keyword vector model* that is commonplace in Information Retrieval. The main variations in the model are mainly in what words serve as keywords, how to weight those keywords and how to compute the similarity functions between two vectors.

3.1 Results

Our tests have shown that the best keyword extraction technique is to use full-text indexing. The best weighting function is the famous « Term Frequency * Inverse Document Frequency » which gives a bigger weight to keywords that are rare in the whole set of texts and frequent in a single text. As for the similarity function it is the cosine function, which gives 1 for similar vectors, and 0 for totally different ones.

After the computation of similarity coefficients and retaining the best 1000 links, we have observed a precision rate of 80%, which means that 80% of the generated links have been judged 'relevant'. The recall rate has been 25%, which means that 25% of existing manual cross-references have been retrieved. This result might imply that our techniques are not very efficient, but the high precision rate means that even if we did not find back many manual cross-references, what we had generated was still meaningful.

More sophisticated keyword extraction techniques could be considered. For instance, identifying noun pairs or concepts. But this requires either important domain-dependent knowledge bases or natural language analysis, which is beyond the scope of this paper.

3.2. Using relevance feedback

In Probabilistic Information Retrieval (PIR), when some knowledge is available about the relevance of a document with regards to a query, the keywords can be re-weighted according to this information, so as to maximise the similarity function for the relevant pairs query/document. The reader interested in PIR will find more information in [22, 2].

Thus, the 701 existing cross-reference links have been used to weight the 4096 keywords that had been identified in the text. We have noticed that there was a major improvement in the recall rate when we used this sort of *relevance feedback mechanism*. More existing cross-references were retrieved (30 %), and so the quality of other generated links proved superior to other methods. This has not been fully confirmed as a precision of 82 % had been observed when browsing through the generated links (about 400 out of 700). Yet this result was interesting, as it maximised the retrieval of links whose type had been judged relevant by somebody.

3.3. A semi-automatic procedure

After the experiment carried out above, we have outlined a semi-automatic procedure for cross-referencing in a hypertext. Its first step could be to run an automatic linking algorithm on the set of documents to link (with simple full text indexing and TF * IDF similarity computation).

The second step would require some human relevance feedback about whether the links generated during the first step are relevant or not. In our case, for example, we have been able to evaluate 400 links out of 700 generated. The more, the better of course.

This could lead to a second running of the linking algorithm where better links could be discovered and implemented automatically.

This opens the path to semi-automated cross-referencing link generation. One will notice that this procedure is very similar to the one used in Information Retrieval for relevance feedback. As its efficiency has been acknowledged many times in that field, this seems to be the case here too.

4. Disability guidelines

After gaining initial experience with the process of hypertext conversion and cross-references generation on a general HCI guidelines book, we have processed other sets of HCI guidelines. Smaller sets of guidelines are indeed available on the WWW for those wishing to make their software more accessible to the disabled. We have thus generated on-line hypertext guidelines from the following sources: [17], [9], [19] and [5]. The result of our hypertext conversion process can be consulted on the WWW (http://www.info.fundp.ac.be/httpdocs/ guidelines/).

The first step of the process has been the generation of a hierarchy of HTML files, all related to one another by hierarchical links. Then we have had to address two separate issues. The first one is the cross-referencing within a single set of guidelines (generation of internal links). The second one was the generation of cross-references between different sets of guidelines (generation of external links).

4.1. The hierarchical links

The starting point for all the conversions into hypertext has been ASCII files containing all the guidelines (one file per source). We have then added some structure to them thanks to a word processor (separating headings of level 1 from level 2, and so on). The hierarchy we have introduced by simply attaching styles to paragraphs was the basis for the generation of a well-structured set of HTML files, each containing a single guideline, and related to one another by simple navigational links (next, previous, top, bottom). This job was achieved by a simple RTF to HTML converter.

The main advantage now is that all these sets of guidelines have the same presentation on the WWW. We have introduced consistency in the presentation of all these guidelines from different sources. From now on, we can easily add new sets of guidelines to our database, from simple ASCII files. One can appreciate the result in *Figure 1*. It is derived from the guide [9] that we have converted into hypertext.

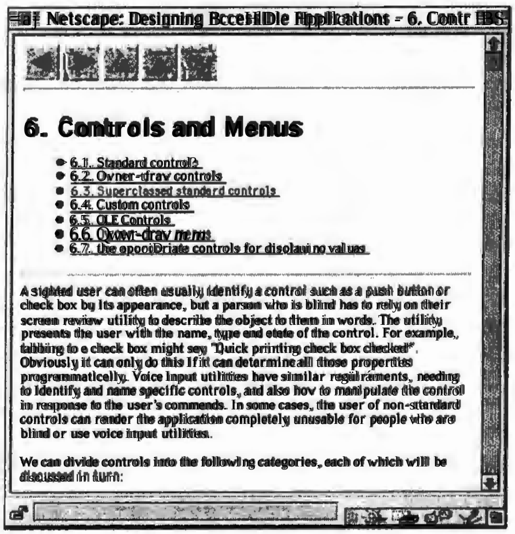


Figure 1 : HTML file with hierarchical links

4.2. The internal links

After the generation of HTML files, we have used our algorithms to produce the internal cross-references within given sets of guidelines. A matrix of similarity coefficients between all the guidelines has been computed. And we have kept the best 1% of them to automatically generate cross-reference links. This has produced good results, though, given the relatively small size of each guidelines set, the precision rate is not really significant here.

4.3. The external links

We have also computed similarity coefficients between guidelines belonging to different sets. The idea was also to retain the highest similarity coefficients to generate external cross-references.

4.4. Evaluation

The results of the automatic cross-reference links generation are summed up in *Figure 2*. The four sets of guidelines are in rectangles. The internal links are represented by semi-circles, while the external links are plain lines. Lines and circles are labelled with numbers giving the number of links [relevant/not relevant] generated by our technique. The thickness of a line/circle grows with the number of links. There is a total of 84 links (either internal or external), 66 of which have been judged relevant (by ourselves), and 18 not. This is an interesting result, as it shows that we have once again reached a precision rate of 80% (this is in fact the *relevance rate*, if you are not acquainted with the technical terms).

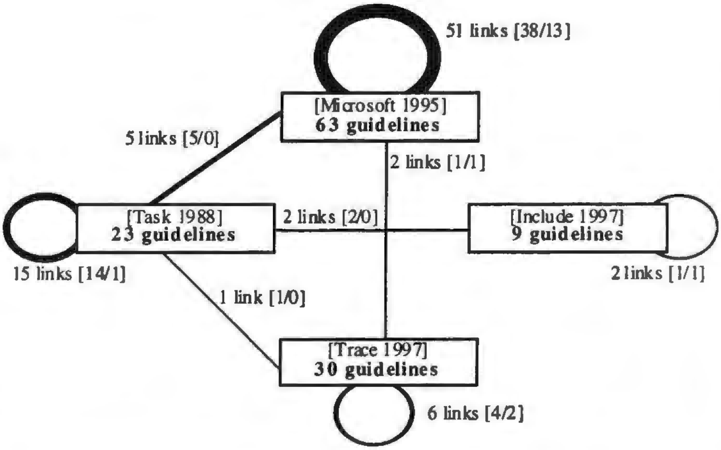


Figure 2. The generated links, both internal and external to our 4 sets of guidelines.

There are quite a few things we can learn from this graphic. First, it shows that most of the generated links are internal ones. It might be because people tend to define and use their own vocabulary when they write a guide, thus sharing fewer common words with other people. Or it might also be that the guides merely do not address the same issues. This is particularly true with the Public Access Terminal Guide [5], which deals a lot with hardware devices such as cards, that are not covered by other guides. This indicates that external cross-referencing is much harder than internal cross-referencing.

However, you may take a look at a good external link in *Figure 3*. It shows a good example of an automatically generated external link. The similarity coefficient between the two guidelines is 0.16 and one can notice that these guidelines are very close to each other indeed.

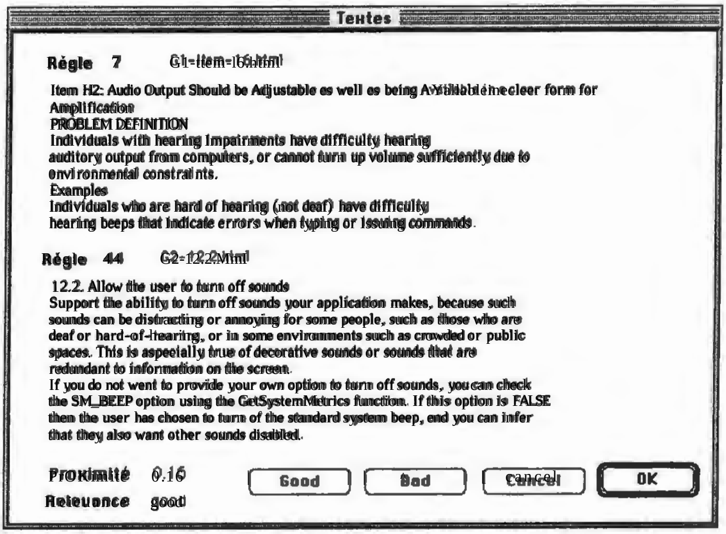


Figure 3. A good external link between two guidelines belonging to different sets.

The other major problem is the 18 links generated by our algorithm that we have rejected as they were not relevant. Here, the main fact is that statistical measures alone do not access the meaning of sentences, and thus produce cross-references that are not always relevant. The meaning of a word is of course not "context free". In *Figure 4*, there is a sample of a bad external link produced by our algorithm. Although the similarity coefficient is 0.15, one can notice that these guidelines do not address the same topics. The confusion here arises from the use of the same words in different contexts. The word 'control' has (at least) two different meanings, thus bringing confusion in the process of similarity detection. That reveals why a purely automatic algorithm will always yield some bad cross-references that should be avoided in the final hypertext version of HCI guidelines.

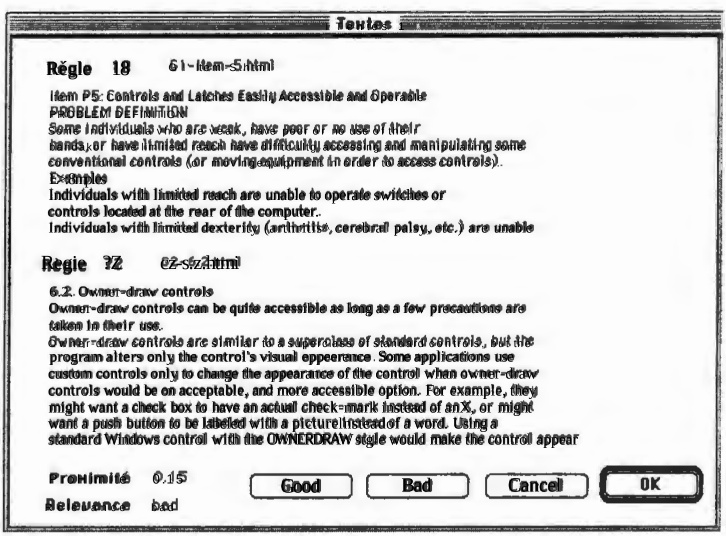


Figure 4. A bad external link between two guidelines belonging to different sets.

5. Conclusions

The statistical methods alone cannot generate well-linked hypertexts, but they can provide valuable help in the framework of a semi-automated linking process. After an automated first step (the computation of similarity coefficients based on keyword frequencies), relevance feedback can be given to the process to yield better cross-reference links.

A human expert is thus needed. His role would also be to organise the guidelines before the conversion process, so that they are neither too long, nor too short, and can be integrated in an existing on-line cross-referenced hypertext database.

Anyway, we feel that an hypertext database on the WWW, with good cross-reference links is a valuable tool to software designers that try to include accessibility issues in their work. Our work might help them if we manage to gather further guidelines to feed our database (if you have some guidelines available, please contact us). And it is accessible from all over the world thanks to the Internet.

6. References

[1] CONRAD, J. G., A System for Discovering Relationships by Feature Extraction from Text Databases. CIIR, University of Massachussets, Amherts, USA, 1994.

[2] FUHR, N., BUCKLEY, C., A Probabilistic Learning Approach for Document Indexing. ACM transactions on Information Systems, Vol 9, N°3, pages 223-248, July 1991.

[3] FURUTA, R., PLAISANT, C., SHNEIDERMAN B., Automatically transforming regularly structured linear documents into Hypertext. Electronic Publishing, Vol 2, N°4, 211-229, December 1989.

[4] GOFFINET, L., NOIRHOMME, M., Automatic Hypertext Link Generation based on Similarity Measures between Documents, RP-96-034, Institut d'Informatique, 1996. http://www.info.fundp.ac.be/~lgo/

[5] GILL, J., Access Prohibited ? Information for Designers of Public Access Terminals, http://www.eyecue.co.uk/pals/

[6] KWOK, K. L., A Network Approach to Probabilistic Information Retrieval. ACM Transactions on Information Systems, Vol 13, N°3, July 1995.

[7] LEHNERT, W. G., Automating the Construction of a Hypertext System for Scientific Literature. Proceedings of the AAAI Workshop on Communicating Scientific and Technical Knowledge, 1992.

[8] MCCANN, A., Designing accessible learning materials for learners with disabilities and learning difficulties, Workshop on Guidelines for the Design of HCI for People with Disabilities, INTERACT'97, Sydney, July 1997.

[9] Microsoft Corporation, Designing Accessible Applications, Accessibility and Disabilities Group One, Microsoft way, Redmond, WA 98052-6399, USA, July 13th 1995

[10] US Nuclear Regulatory Commission, USA.

[11] HECTOR Chris (chris@sunpack.com)

[12] SALMINEN, A., TAGUE-SUTCLIFFE, J., McCELLAN, C., From Text to Hypertext by Indexing. ACM Transactions on Information Systems, Vol 13, N°1, January 1995.

[13] SALTON, G., Automatic Text Processing. ISBN 0-201-12227-8. Addison-Wesley, Reading, Massachussets, USA, 1989.

(14]SARRE, F., GÜNTZER, U., Automatic Transformation of linear Text into Hypertext. Proceedings of DASFAA'91, pp 498-506, Tokyo, 1991.

[[13]SMITH, S., MOSIER, J., Guidelines for Designing User Interface Software. ESD-TR-86-278. MITRE Oitporation, Bedford, Massachussets, USA, August 1986.

' [léjSTEFANIDIS, C., Designing for Accessibility : Reflections on the Use of Human Factors Design Guidelines, Workshop on Guidelines for the Design of HCI for People with Disabilities, July 1997, INTERACT'97, Sydney.

I [17JConsiderations in the Design of Computers to Increase Their Accessibility by Persons with Disabilities, tadustry/Government Cooperative Initiative on Computer Accessibility Task Force, Trace Center, S-151 Waisman Cnter, 1500 Highland avenue, Madison, W1 53705, USA. http://sjuvm.stjohns.edu/rehab/tech94/aging.html

(1i)^OREN, C., Nordic Guidelines for Computer Accessibility, Nordic Committee on Disability, Box 510, S-162 15, Vallingby, Sweden.

j 19]Unified Web Site Accessibility Guideline, ADVANCED DRAFT COPY. Page Author, User Agent and Assistive Tfdinology Guidelines and Issues - Version 8, Trace Center, S-151 Waisman Center, 1500 Highland avenue, (iWison, W1 53705, USA. http://trace.wisc.edu/

S0]UENO, K., OGAWA, K., A Design Guideline Search Method that Uses a Neural Network. HCI'93, Vol 19B.

PIJVANDERDONCKT, J., Accessing Guidelines Information with SIERRA, Proceedings of Interact'95, p 311-316. j^Qupman & Hall, 1995. Available at http://www.info.fundp.ac.be/~jvd/soft.html

PIJVAN RIJSBERGEN, C. J., Information Retrieval Butterworths, London, England, 1979.

SPECIALISED KEYBOARD FOR WHEELCHAIR-BASED COMPUTER ACCESS

Van Well, J.P.M., Cremers, G.B., Koomen, G.

Abstract

This paper reports on the development and usability evaluation of a specialised keyboard, in early prototype phase, for wheelchair-based computer access. The objective has been to obtain early responses on the functional and technical aspects of keyboard prototype design from expert users. Typing performance was measured and compared with that of a standard keyboard. Recommendations were made for improvement of the layout of the prototype keyboard. This evaluation has confirmed that the application of a miniature keyboard for wheelchair-based computer users is a promising concept for people with reduced manual dexterity.

1. Introduction

Standard QWERTY keyboards used on most personal computers are often inaccessible to people with impaired mobility and reduced manual dexterity. Many disabled users do not have the strength required to press the keys on a standard keyboard. Those people with limited range of motion are not able to move their hands easily from the alphabetic keypad to the arrow keys, function keys and numeric keypad. Others with uncontrolled or involuntary hand movements make frequent typing errors by pressing the wrong key, or by pressing a key longer than normal, inadvertently activating the automatic key-repeat feature of many keyboards.

A miniature keypad and a keyboard emulator may solve some of these problems, especially that for people with limited range of motion, but unaffected fine motor control. Considering keyboard usage while seated in a wheelchair gives a second reason for selecting a mini-sized specialised keyboard. Big keyboards take up too much space. Mobility is under this condition a main requirement, whereas typing is of secondary importance.

For a wheelchair demonstration platform being developed under the ACTS Umptidumpti project a first prototype of such a miniature keypad with emulator for personal computing was built. This prototype has been evaluated with the help of able-bodied subjects. The prototype keyboard, its evaluation and results are described hereafter.

2. Goal

The evaluation is carried out to provide initial quantitative and qualitative responses from able-bodied users on the functionality, ease of use, performance and overall opinion of the keyboard prototype. The results will be used to optimise software characteristics of the keyboard emulator and input devices. This will be done with respect to the following aspects: keypad layout optimisation, keypad ergonomics, system feedback, timing behaviour of key press and release. In a later stadium, aspects as functionality to allow switching between input devices or properties for simultaneous and multiple input device usage will be evaluated.

3. Method

Keyboard input with the keypad and the keyboard emulator has been tested for alphabetic typing. Typing speed between a standard laptop keyboard and the miniature M3S keypad has been compared. Subject's opinion on the overall system has been recorded. Two additional layouts for numeric keying and cursor key navigation were briefly tested. A rough indication on learning effects is measured through repetitive attempts on the same typing task.

4. Subjects

The subjects were eight able-bodied male and female researchers, in age between 20 and 58 years. Two of them were expert users (designers) of the keyboard emulator system. The other six subjects had never used the keypad system before. All subjects had a higher or university education and extensive personal computer experience.

5. Experimental set-up

Subjects were given the task to type an 11 word (72 character) sentence in English language with both keyboards. Task completion times were measured. The typing task was carried out on a portable computer with a standard QWERTY keyboard (5), Microsoft Windows 95 operating system and Notepad text editor. The task completion times were clocked with a stopwatch. The task was then repeated with the mini keypad. In figure 1, a block diagram of the experimental set-up is shown. A 16 key mini keypad (1), especially made for use on a wheelchair, is connected to a keyboard emulator (3) via an M3S-wheelchair bus-system (2). A description of M3S, the Multiple

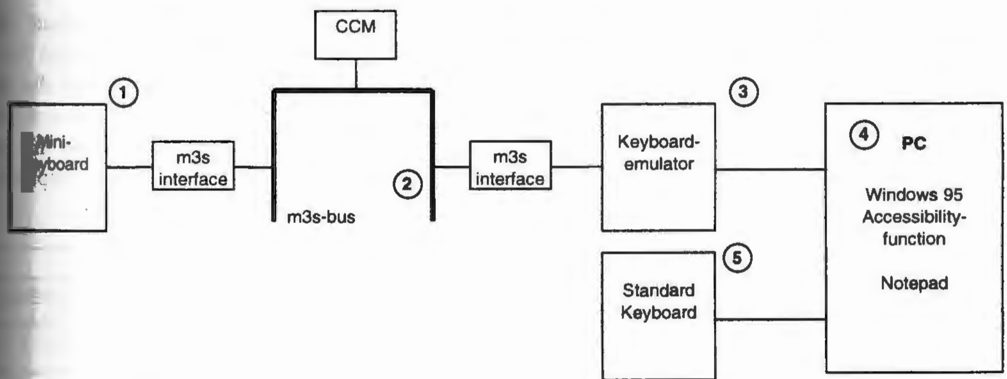


Figure 1. *Block diagram of the M3S miniature keypad and keyboard emulator evaluation*

Master Multiple Slave bus system, is given by Nelisse, M.W., Overboom, GR. et al [9] and in the Uffiptdumpti deliverables 10 and 13 about Functional [12] and Technical Specifications of Applications and Services [11]. Brief information is also available on the internet at http://www.tno.nl/m3s. The keyboard emulator receives 16 on/off signals of the keypad, translates these into QWERTY keyboard commands and passes them onto the portable computer (4) via a serial bus at 19 Kbytes/sec. Inside the personal computer these commands are interpreted by the "Serial Keys" feature of the standard Windows 95 "Accessibility" function and passed onto the Notepad text editor.

The typing on the mini keypad was done according to the layouts as shown in figure 2. For the sentence typing task only the alphabetic layout (left) was used. The key selection works similar to alphabetic keying on a telephone keypad. Pressing the key with the character of your choice repeatedly until it appears on the computers' screen, then acknowledge this selection with the bottom left OK key and move on the next character. For corrections, one uses the C-key, which behaves as the Backspace-key on a QWERTY keyboard. The Enter key has the same function as on the QWERTY keyboard. The black > key arrows the user to switch onto the next layout. These are used in similar fashion for numeric and cursor navigational keying. Note that on the latter two layouts, the use of an OK key for acknowledgement is not needed, because there are no multipl function keys.

C	Abc	Def	ghi
	Jkl	Mno	pqr
	stu	Vwx	yz
Ok	➤		Enter

C	1	2	3
	4	5	6
	7	8	9
	➤	0	Enter

C		⇧ IF	
Esc	⇦		⇨
Tab		⇩	
	➤	Alt	Enter

Figure 2. *Diagrams of MSS miniature keypad layouts. From left to right: alphabetic, numeric, and navigational la*

6. Results & Discussion

6.1. Quantitative Data Analysis

In figure 3 and table 1 the time to task completion for eight different subjects is visualised. The mean task completion time for the QWERTY keyboard and the mini-keypad were 40,2 and 360,[illegible] seconds respectively, with standard deviations of 15,2 and 63,1 respectively. This means that the average typing performance with able-bodied subjects is approximately 8,9 times faster with the QWERTY keyboard compared with the prototype mini keypad. However, it is expected that for motor impaired subjects, who are much slower on QWERTY keyboards can reach a compa[illegible] speed with both keyboard types.

Task completion [in seconds]								
Input Device	***Subject 1***	***Subject 2***	***Subject 3***	***Subject 4***	***Subject 5***	***Subject 6***	***Subject 7***	[illegible]
Normal keyboard	28.0	26.8	36.0	74.0	48.8	42.0	22.8	[illegible]
Miniature keypad with emulator	495.0	360.0	391.0	396.0	316.0	315.0	277.0	336.2

Table 1. Time to task completion for the simple alphabetic typing task using two different keyboards

The small size of a mini keypad reduces the need for problematic gross motor activity on the elbow and shoulder and can therefore result in a better typing performance.

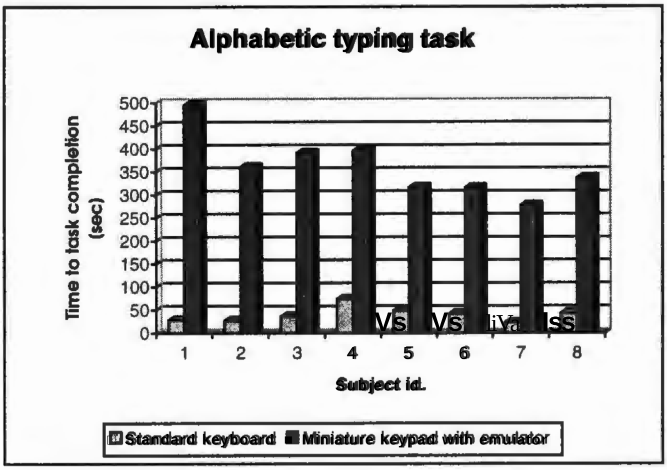


Figure 3. *Time to task completion for simple alphabetic typing task using 2 different keyboards*

Task completion [in seconds]				
Input Device / Subject 8	***1st Session***	***2nd Session***	***3rd Session***	***4th Session***
Normal keyboard	43,3	31,0	25,6	23.0
Miniature keypad with emulator	366.2	276.7	235.2	233.2

Table 2. *Time to task completion for the four session repetitive alphabetic typing task*

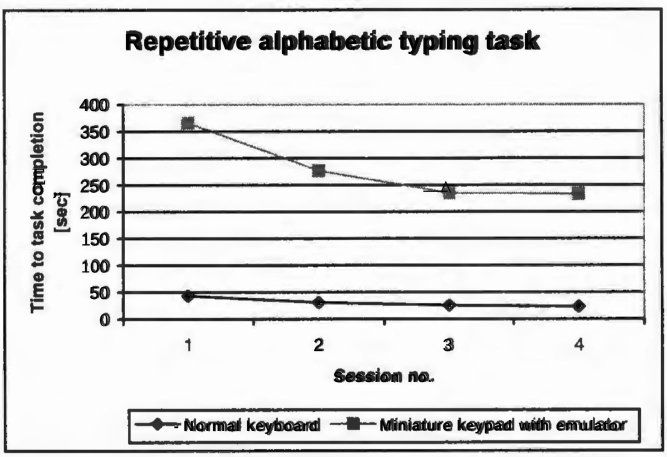


Figure 4. *Time to task completion for repetitive alphabetic typing task using 2 different keyboards*

Figure 4 and table 2 show the time to task completion scores for the repetitive alphabetic typing task. At the fourth session with the normal QWERTY keyboard, the subject was able to bring back his time to 53 % of the time he needed in the first attempt. On the mini keypad the time reduced in a similar way to 64 %. This may mean that the subject learned to type sentence by heart, or that he got used to use the keyboards faster, or just tried harder to be faster. However, the order of magnitude of speeding up through learning is roughly the same for both keyboards. Nonetheless, it is much too early to draw conclusions from these few sessions. Additional tests are necessary with more subjects, more sessions and a variety in typing tasks.

6.2. Qualitative Data Analysis

6.2.1. Using the alphabetic layout

In some cases, the system was not fast enough to capture rapid repetition of keystrokes. This slowed down the user during text entry. The reason for this probably lies in the performance of the Windows 95 accessibility function. The longer the escape sequence it needs to interpret is, the longer it takes the computer to carry it out.

In some cases, the abc-key did not start with an 'a' but with a 'b' instead. This was different from the user expectations of keyboard emulator software functionality. The result was that the users made many typing errors and more importantly lost a lot of time recovering from these errors.

The keys of the layouts need to be clearly indicated on the keys of the keypad themselves. This demand may get less important after users are trained with the keypad. Then they will know where the specific keys are located on the pad. After some time of training "you will get a feel for it", one subject said.
Other subjects liked to have a visual separation between the block of nine keys for a to z and the remaining seven control keys for correction, OK and Enter.
One subject liked to have a little knob on the mno- key, similar to the knobs on the f a j key ons QWERTY keyboard, so that you can feel that your finger is in the starting position.

An Alt key could also be added to the alphabetic layout, to enable shortcuts like alt-f for file or ali-t for exit. This would however only work if alt and f could be pressed at the same time or with a feature that allows to press alt and f after each other.

The alphabetic layout was further far from complete, e.g. capitals-, diacritics- and keys for special signs need to be added.

6.2.2. Using the numeric layout

The keying on the numeric layout was experienced to be straightforward. The fact that no OK key needed to be pressed, was very positive to the user. Negative points however, were ck incompleteness of the numeric keypad layout: decimal point, colon, plus, minus, asterisk, divide were lacking as well as auto repeat key features.
Very confusing was the fact that it was never clear, which of the three was the active layout. Clei! feedback with lights, or pro-prioceptive feedback through a separate and unique key press per layod selection, are required.

6.2.3. Using the navigational layout

The navigational layout could only be used after the special key Fn F10 was pressed on the portabkJ computers' keyboard. This enabled the cursor key mode on the computer. On desktop PC's wi I for 103 key keyboards, this is not necessary. Nevertheless, this caused a lot of confusion. Once this was mastered pulldown menus could be operated with relative ease.

7. Conclusions

The keypad prototype implements a usable way of key selection, and is easy to learn. The average typing performance, with able-bodied subjects, reached in this test is 11.2 % of that of a normal QWERTY keyboard. Subsequent tests with motor impaired will need to be carried out. We expect a far higher percentage in that case, as the small size of the keypad minimises the amount of required motor action. The technical performance of the keyboard emulator was not in all cases sufficient. In certain conditions the keying process was slowed down by the fact the characters just did not appear fast enough on the screen.

The multiple layout approach of the keypad system clearly needs feedback, to indicate which layout is the active one. Layout(s) need to be clearly indicated on the keys of the keypad. The layouts need to be completed, so that all possible PC keystrokes can be emulated.

Further usability tests are necessary to optimise the individual layouts, in such a way that the maximum number of keystrokes can be realised with a minimum amount of motor action on behalf of the user.

8. Acknowledgements

The author acknowledges the valuable support from Mr. R.H. Disco, student Hogeschool Heerlen and Mr. M. Nelisse, M3S Dissemination Office, TNO Institute of Applied Physics during design and technical implementation of the prototype device.

9. References

[1] ARNOTT, J.L. et al., Small Text Corpora in Character Disambiguation for Reduced Typing Keyboards, Proceedings of the 13th RESNA Annual Conference, Washington DC., 1990.

[2] FOULDS, R. et al., Lexical Prediction Techniques applied to reduce Motor Requirements for Augmentative Communication, Proceedings of the 10th RESNA Annual Conference, San Jose, California, 1987.

[3] HINE, N.A., BEATTIE, W. AND ARNOTT, J.L., Factors Affecting Usability of Telecommunications Services by People with Disabilities., Proceedings of 16th International Symposium on Human Factors in Telecommunications May 1997, pp 201 -208.

[4] HELANDER, M., Handbook of Human-Computer Interaction, Elsevier Science Publishing Company, Inc. New York, 1998.

[5] MELE I. VAN, EDWARDS K., VERHEUST M., SPAEPEN, Mobile access to WWW services for university students who are motor impaired: a qualitative study on man/machine interface, Human Factors in Telecommunications '97 Oslo, Norway.

[6] ANONYMOUS, Customising Microsoft Windows 95 for Individuals with Disabilities, Microsoft Press, 1995.

[7] LINNEMAN, S.L., "A Simplified Touch-Tone Telecommunication Aid for Deaf and Hearing Impaired Individuals", Proceedings of the 8th RESNA Annual Conference , Memphis Tennissee, 1985.

[8] MUNRO, A., Functional Requirements for Special Needs, ACTS Mobile Summit 1997 Proceedings, Vol. 1. Aalborg, DK, 1997.

[9] NELISSE, M.W., OVERBOOM, G. et al., A general-purpose integrated and modular architecture for the rehabilitation environment; Mobile Robotics Technology for Health Care Services; Proceedings of the 1st MobiNet Symposium, Athens, GR, 1997, pp. 161-170.

(10) RUBIN, J., Handbook of Usability Testing, John Wiley & Sons, Inc. New York, 1994, pp. 25-46.

[11] VAAN, J. DE, et al. Technical Specifications of Services and Applications, Umptidumpti deliverable no. 13, ref: AC027/UMPTI/IvD/DS/I/D13/b1, 1997.

[12] WELL, JPM. VAN, et al., Functional Specifications of Services and Applications, Umptidumpti deliverable no. 10, ref: AC027/UMPTI/IRV/DS/I/D10/b1, 1997.

HUMAN-COMPUTER INTERFACE BASED ON COMPUTER VISION

Joan Aranda, Joan Climent and Antoni Grau(*)

Abstract

A new man-machine interface has been developed based on computer vision that permits real time control of cursor position by the glance. The proposed interface consists of a camera connected to a PC fitted with a specific low cost hardware image processor implemented for this application. This image processor performs at video rate a discrete polar transformation on contour image that minimises computational cost of the motion quantification algorithm executed by PC processor. Local motion is estimated basically by subtracting the data provided by image processor relative to consecutive frames. Global motion quantification in the image is obtained from a set of these local measures over the image. The presented interface reduces physical barriers to people with several disabilities (as tetraplegic people) permitting them to access to computer technology and all its applications.

1. Introduction

In this paper a new human-computer interface is presented to move the cursor position over the screen. Keyboards and mechanical mouse over the table present an access barrier to computer technology to some people with some kind of severely physical disabilities (as tetraplegic people). They limit the access to powerful applications to increase his/her personal development and degree of autonomy. Among these applications we can show up educational interactive software, peripheral devices control (for instance at home) and communication capabilities (as access to internet).

(*) Dep. of Automatic Control and Computer Engineering (UPC).

Some efforts have be done in the design of specific mechanical interfaces for handicapped people. However, mechanical interfaces are sometimes annoying and moreover some of them require for a second person for installation. On the other hand, speak recognition have been successfully used in [4] and [6] as a promising interface for high level oral orders, but it seems not suited for accurate control of cursor position.

To overcome these inconveniences, a new interface has been developed based on computer vision. The architecture of presented interface is based on a PC fitted with a specific image processing board. The camera is connected to the PC through this specific image processor. Then, user movements in front of the camera are detected and measured in order to be used as input to PC mouse driver. In this way, user can move cursor over the PC screen by moving his/her head or any part of his/her body.

The image processor board has been implemented in our laboratory and is a low cost hardware based on programmable components that can be adjusted to application. A PC has been chosen as a host for this specific processor also due to its low cost.

Different methods can be used to detect and measure present motion in the visual field of a camera such as image difference, Fourier Transform and optical flow. Also tracking of segmented objects inside the image can be used. In [5] a human computer visual interface is presented consisting on a wireless glove with four coloured markers. The user is wanted to put this specific glove on and move it in front of the camera to control a 3D graphic CAD model on the computer screen. The system presented in [1] is able to recognise and track a predefined feature in the user face (corresponding to the middle point between the eyes). In [2] a new tracking method is presented that can follow any predefined singular point in the image.

As proposed in articles [1] and [2], the presented interface doesn't require the use of hands neither the use of specific markers sited on user body. The proposed method to compute the direction and quantity of the present motion in the image, is based on the analysis of a set of differential local measurements over the image (or over a predefined window in the image). Local motion detection and quantification are obtained from an incremental procedure based on a specific image transformation of contour pixel coordinates.

2, System description

With the aim of detect and quantify the motion generated by the user we have developed a new visual interface consisting of a camera connected to a PC fitted with a low cost specific image processor. This image processing board provides to PC only the relevant information for motion quantification. This information is extracted at video rate from the input image in two steps, first contour extraction is performed and second, a specific image polar transformation is applied. The complete architecture of the image processor will be presented in section 4.

The acquired image is divided up to 32x32 processing windows of 15x15 pixels. These windows are overlapped in 7 pixels (almost half window) in the horizontal direction and also in vertical direction, so that a 263x263 region of the image is processed. As can be seen, central windows of this region are fully overlapped. Only the external 7 pixels frame around this region gives non overlapped information. This 32x32 grid of overlapped windows provides with redundant information in order to increase the robustness of the motion quantification algorithm.

Contour pixels coordinates included in these processing windows are transformed so that their relative position from central pixel of window is calculated. This computation is performed based on a discrete polar transformation of the image. We use a polar coding with less resolution than others authors with the aim to attain video rate coding with a low cost hardware image processor. In this way, only eight distance values from central point of the window are obtained, corresponding to the distances in the eight main directions (N, NE, E, SE, S, SW, W, NW). This defines a distance vector $p(\theta)$ $\theta \in [0..7]$ that codifies the position of contour pixel inside every processing window (see *figure 1*).

3. Proposed algorithm

For every processing window in the image, a distance vector $\rho(\theta)$ is acquired by the PC at video rate. This means that 1K distance vectors $\rho(\theta)$ are acquired every 20 ms. These distance vectors determine the position of contour pixels at every frame in a proper format. From these raw data, local motion can be evaluated by subtracting the distance vectors of two consecutive frames.

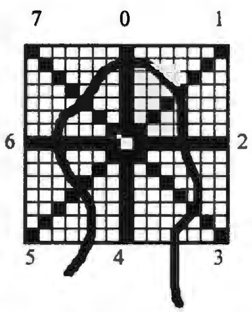


Figure 1. a) Distribution of radii in a processing window

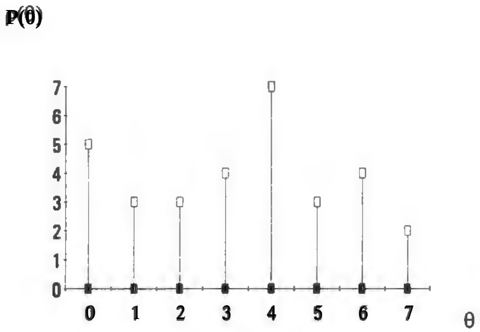


Figure 1. b) Resulting distance vector p(θ) for the example in figure 1 a).

Let $p_{t-1}(\theta)$ and $p_t(\theta)$ be the consecutive distance vectors of the same processing window in the image, corresponding to frames *t-1* and *t*. Then we can define the incremental distance vector $d_t(\theta)$ as:

$$d_t(i) = p_t(i) - p_{t-1}(i) \quad \forall i \in [0..7]$$

From this distance vector, Δx and Δy of contour pixels can be locally estimated for every processing window by using the next formulae:

$$\Delta x = d_t(2) + d_t(1).\cos(45^\circ) + d_t(3).\cos(45^\circ) - d_t(6) - d_t(5).\cos(45^\circ) - d_t(7).\cos(45^\circ)$$

$$\Delta y = d_t(4) + d_t(3).\cos(45^\circ) + d_t(5).\cos(45^\circ) - d_t(0) - d_t(7).\cos(45^\circ) - d_t(1).\cos(45^\circ)$$

Since these computations have to be done for every processing window (so 1K times per frame) a look up table (LUT) has been used to optimise computation time. These 32x32 $(\Delta x, \Delta y)_{ij}$ values constitute the raw information for global motion estimation present in the image. At every new frame, average global motion components $(\Delta X, \Delta Y)$ can be obtained from geometric mean of local motion data:

$$\Delta X = \frac{\sum_{i,j} \Delta x}{1024} \qquad \Delta Y = \frac{\sum_{i,j} \Delta y}{1024}$$

A temporal filter has been finally included in order to increase stability in front of eventual error in global motion estimation (like those produce by light fluctuations that can momentary disturb the contour extraction process). The selected filter consists of a median filtering of last five avera global motion estimations (AX, AY). That is, the most probable value of the last 1/10 sc. This kind of filtering has given better results than classic lineal temporal filtering. The quantity of frames to take into account can be adjusted by the user himself by means of an application menu.

Finally, the filtered global motion estimation is used to actualise the cursor position in the screen by incrementing the actual position. The (x, y) increment of cursor coordinates can be adjusted to application with the aim of guarantee comfort in user movements and accessibility to the whole screen.

An easy strategy have been applied in order to simulate the click button of the classic mechanic mouse to "open" a software application. This permits to an handicapped user the access to most of actual icon based operating systems and computer applications. To select one application, the user only has to keep the cursor over the corresponding icon for few seconds. The exact time is obviously adjustable by the user.

4. Specific hardware

In order to get video rate control of cursor position without consume most part of PC processing capabilities, a specific low cost image processor has been implemented in our laboratory. This processor feeds to PC at video rate (50Hz) with data from polar transformed contour image, that is, with the IK distance vectors ($\rho(\theta)$) of corresponding processing windows in the image. As explained before, these vectors constitute the raw data for image motion quantification algorithm and cursor movement. *Figure 2* shows the complete architecture of image processor.

In a first stage image processor performs a contour extraction reducing the digitised input image to a binary contour image. This reduces the amount of data to be processed maintaining the relevant motion information contained in the image. Also contour image has shown to be more stable in front of lighting changes. This information is obtained at pixel rate and only pixel clock is needed to control the overall process. Pixel clock has been adjusted to 7,37 MHz in order to work with square pixel shape. This sets the image resolution to 383x287 pixels/frame.

Sequence to parallel conversion is necessary to feed 15 pixels column to the polar transformation module at every pixel clock. This column is fed into 15 shift registers of 15 pixels length, that form a 15x15 matrix (processing window). From this window of 225 pixels only 56 are selected corresponding to 8 radii of 7 pixels length *(figure 1)*. Each of these radii is conducted to a priority coder so the distance of pixel contour in the analysed direction is coded in 3 bits (from values 0 to 6). The value 7 is reserved for the case that no pixel contour would be found in that direction. Polar transformation module has been totally implemented into one FPGA chip, which permits to change the applied image transformation for others applications. More details about the implementation of this module can be found in [3].

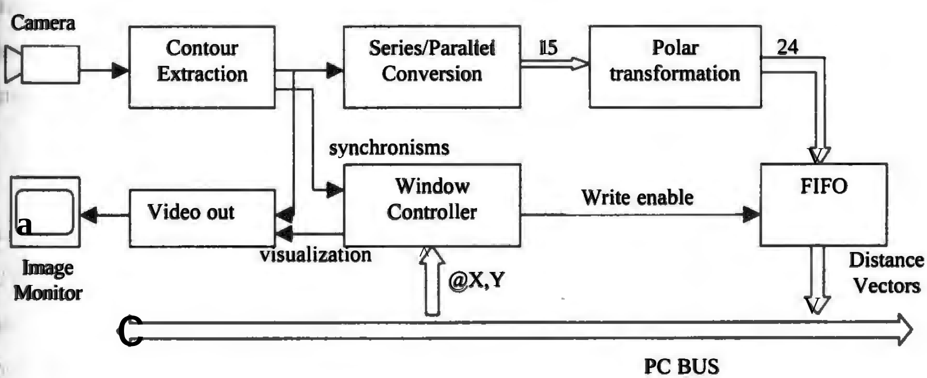


Figure 2. Architecture of the implemented low cost image processor

At pixel rate, 24 (8x3) bits are obtained from polar transformation module which constitutes the distance vector ρ(θ) associated to every pixel in the image. User can program the window controller module to write only the selected information in the FIFOs. For this application, the image processor has been programmed to save only the distance vectors corresponding to a 263x263 central region in the image. Distance vectors information is then only saved every 8 pixels in line and every 8 pixels in column, configuring a 32x32 grid of distance vectors.

FIFO module has been selected as a low cost alternative to dual port memory access. It also acts as a buffer permitting the PC processor to acquire the selected distance vectors of last frame while image processor writes those relative to present frame. So the read and write processes are overlapped in time, obtaining a sampling frequency of 20 ms.

5. Results

A robot test-bed has been designed before proving the interface with human users. For this experiment, a robot has been located in front of the camera. Then, it has been used to move a set of defined textured visual patterns following pre-programmed 2D trajectories. The selected trajectoria, for the experiment have been a straight line, a square, a rhombus, a star, a sinus and a circle. A robot has been chosen so that repeatability of test could be guaranteed for the same pattern and for different ones. These trajectories have been executed by the robot at different speeds, obtaining good results even when robot moves at higher speed than a PC human user normally do.

Finally, the interface have been tested with human users. We use a standard monochrome camera sited on the PC monitor and pointing to the PC user position (in front of the PC monitor). This location of the camera provides an easy visual feedback of cursor position to the user. Also user can feedback what is in the field of view of the camera through image monitor. From PC monitor, camera has been located so that it points towards user face. In this way, user can move the cursor position with head pan and tilt movements. Cursor movement increment can be adjusted by the user himself in order to get easy control of its position.

We have also experimented with other possible solutions like following the movements of users eyes. This strategy results more tired to the user than moving the head, because it requires from a high self-control of eyes movements and capability of concentration. However, hand movements and body movements in front of the camera have been successfully proved.

6. References

[1] AMAT, J., CASALS, A. AND LLARIO, V. "Man machine communication by the glance", Cognitiva-J?) Pp. 372-376, Paris. 1987.

[2] AMAT, J., ARANDA, J., AND CASALS, A. "Tracking Capabilities in Computer Vision for Industrial Applications". SPIE. Vol. 33, No. 11. 1993.

ARANDA, J. "Image Processing Card for Real Time Target Tracking". DCIS. Sitges, Catalonia. 1996.

CASALS, A., VILLÀ, R., AND CUFI, X. "Tou an assistant Arm: Design and Control Objectives and Performance". International Conference on Advanced Robots. 1993.

CIPOLLA, R., OKAMOTO, Y., AND KUNO, Y. "Robust Structure from Motion using Motion Parallax". Fourth International Conference on Computer Vision. Berlin, Germany. 1993.

DARIO, P. "A Mobile Robotic unit for the Assistance to the Disabled". Rehabilitation Robotics Newsletter. Vol. 5, No. 1. 1993.

A Head Operated 'Joystick' Using Infrared

Roger Drew, Stephen Pettitt, Paul Blenkhorn & Gareth Evans[1]

Abstract

This paper describes the principle of operation, the design and the evaluation of a low-cost head operated joystick that uses infrared transmitters and receivers to determine the head position of a user and then emulate a Microsoft mouse appropriately. Firstly we introduce the requirements of a device to control mouse cursor position by head movement and then, briefly, summarise approaches that satisfy, or partially satisfy, the requirements. We then introduce the principle of using overlapping infra red beams as a basis of determining head position in a fairly accurate, but low-cost, system. The design of our Joystick is then presented, with particular consideration given to the method of signalling between transmitter and receiver and to the techniques used to remove interference and that provide for improved operation. The results of the evaluations of the joystick are then considered including issues such as operational range (to 10 m and more) and power consumption (continuous operational life of 1300 hours). We conclude with a brief overview of the status of the work.

1. Introduction

For computer users who cannot use a conventional, hand-operated computer mouse, one option is to control the position of the mouse cursor by head movement. A 'head operated mouse' has a number of desirable characteristics. The 'mouse' should work as a standard mouse and require no software to run on the computer. The 'mouse' should work with both standard desktop computers and laptops. The 'mouse' should be unaffected, or at least recover from, alterations in the physical environment, e.g.: changes in the level or spectrum of ambient light; changes to the local electromagnetic field; changes in the distance between the user and the computer; and alterations in the orientation of a user and his/her computer. The connection between the user and the computer system should be 'wireless', in the sense that there should be no physical connection between the user and his/her computer. Ideally the system should not require the user to have to 'wear' any equipment. However, a number of practical approaches, including our own, do have this requirement. If the user has to 'wear' some equipment this should be aesthetically pleasing, small, light and comfortable. It is desirable that the clicking of mouse buttons be integrated with the operational approach. In many cases this is not practical and some other form of mouse button activation either by appropriate switches or by 'dwelling' can be provided. If switches are used, the device should permit these switches to be connected to it and the switch depression mapped to mouse button clicks appropriately. In addition, a 'head operated mouse' will be subject to cost requirements, such as low initial purchase cost and a low cost of ownership.

---

[1] Technology for Disabled People Unit, Department of Computation, UMIST, Manchester, M60 1QD, United Kingdom. Tdpu@cl.co.umist.ac.uk

2 Operating Modes

A mouse cursor can be controlled in two ways, namely as an absolute pointing device or as a relative pointing device. For an absolute pointing device, the position of the device corresponds to the position of the cursor. For example, if the mouse cursor is in the centre of the screen and it is desired that the cursor be moved to the top left, the device is moved up and to the left and held in that position until further movement is required. Conventional hand operated computer mice and trackerballs operate in this way. Although in the case of the former, there is the opportunity to pick up the mouse and reposition it on a desktop without altering the mouse cursor's position.

Relative pointing devices work by displacing the device until the desired position of the mouse cursor is reached, then the device is moved back into a neutral area (a 'deadband') when movement stops. The mouse pointer will only move when the device is displaced from this deadband. Standard joysticks work in this way.

Head operated mice can use either principle, although absolute positioning requires typically more precise head position control and suffers from lower immunity to interference in its environment. The device that we describe in this paper is commonly used as a relative pointing device. However, it can be enabled to work as an absolute positioning device if desired, but the resolution is severely reduced (see later).

3 Approaches to Head Operated Cursor Control

There are a large number of approaches to the control of cursor position through head movement based on a number of different operating principles. Space constraints preclude a detailed examination here, but a brief summary of approaches is given.

A number of approaches use the analysis of video images to determine head position. Some systems simply use a camera [1], but require significant amounts of processing and are subject to variations in ambient lighting. Other systems constrain the lighting. One commercial system uses infrared and has a reflective target on the user (the HeadMouse™[2] [2]). Another approach uses laser light. In this type of approach laser light can be directed at the head (with the reflected light used to determine head position) or directed at the monitor (by lasers mounted on the head) with a receiver local to the computer determining head position [3]. Other video based systems mount small image sensors close to a user's eye and determining where the user is looking (for example Video Control Systems Infrared-based mouse). A related approach that uses eye position, but is not video-based is the Eagle Eyes [4] system that detects eye position by measuring voltage differentials around the eyes, through skin-mounted sensors.

Other approaches may be based on determining the head's position relative to a field, for example, relative to the Earth's gravitational field [5], the Earth's magnetic fields [5] and local electromagnetic fields (generated by the computer monitor). The first two methods have been developed into prototype systems at UMIST. They are low cost solutions, but require the users to bend their necks toward their shoulders in order to achieve left or right movement of the cursor and, thus, are tiring to use for long periods. Further approaches may use ranging techniques, timing

---

[2] HeadMouse is a trademark of Origin Instruments, Grand Prairie, Texas, USA

signal propagation (for example ultrasonic signals) between the user's head and a receiver situated at the monitor. Other approaches may use gyroscopes and accelerometers to detect movement.

The goal of our work over a number of years has been to produce a low-cost head operated system which is reliable and easy to use. The approaches based on video images are generally quite expensive as processing requirements are high and in many cases sensors and light sources may be expensive. Other approaches, for example those based on accelerometers or gyroscopes have high component costs. For instance the cost of a pair of gyroscopes in small quantities exceeds the cost of the entire system described in this paper by a factor of five. Other approaches, such as those developed at UMIST based on the Earth's magnetic and gravitational field have undesirable operating characteristics. The systems based on detecting eye position and determining where the user is looking may offer low cost alternatives. Indeed the target price of the Vision Control System is £100 [6], although this would appear to be dependent on mass-market quantities driven by the games industry. It also has a requirement that users keep their heads relatively still during use. Our approach is not dissimilar to the head-mounted laser approach described earlier. However, it uses lower cost components (infrared LEDs) and a very simple detection mechanism.

4. Principle of operation

4.1 Basic Principles

An LED could be chosen that operates in the visible spectrum or at infrared frequencies. However, LEDs are more efficient at wavelengths greater than 900nm (nearly infrared) and silicon photodiodes have sensitivity peaks have around 900nm, so infrared gives better range [7].

The irradiation pattern of an LED is not uniform in all directions. Figure 1 shows a polar plot of an idealised irradiation pattern of an LED. Consider point B, which is the same distance from the LED as point A (OA=OB), but is displaced θ°. The radiant intensity here is reduced with respect to A due to the non-uniformity of the LED's irradiation pattern. Now, we can relate, in theory at least, the irradiation at a point to its angular displacement from a given position, or, alternatively, relate the irradiation at a fixed point to the relative angular displacement of an LED.

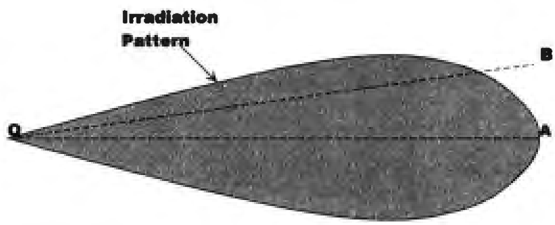


Figure 1: Polar Plot of an idealised irradiation pattern for an LED

A single LED and single receiver are insufficient to determine the relative position in two-dimensional space. Consider Figure 2, which shows contours of equal irradiation at some fixed distance. Point A, the centre of the beam, has the maximum irradiation. Points B, C, D, E all receive the same irradiation, yet are in different positions relative to the maximum. To distinguish between these points we need to use multiple LEDs together with a single receiver or a single LED with multiple receivers. The principle is explained using multiple LEDs and a single receiver, and here we constrain the discussion to determining the position in one dimension, that is, considering Figure 2 whether we are at position B or position C.

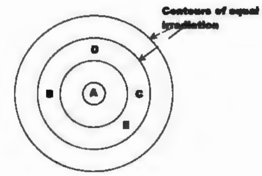


Figure 2: Contours of equal irradiation at some arbitrary distance

Ibbasis of the approach is to use a pair of LEDs to determine angular position. Consider Figure 3. Here we have two LEDs (X and Y) whose idealised irradiation patterns overlap. As one moves Imposition A to position B, the irradiation due to X decreases and Y increases, similarly as one mes from A or C, the irradiation due to X increases and Y decreases. If the light emitted from X IÉY is encoded, so that the receiver can tell them apart, the receiver can determine the irradiation 4rtoeach. With appropriate processing, this method can be used to determine the relative angular position.

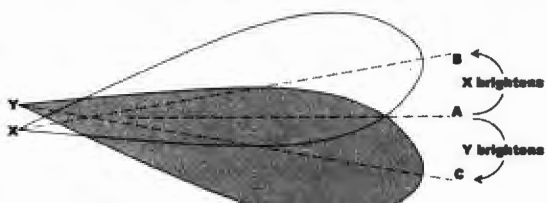


Figure 3: Overlapping Irradiation Patterns from two LEDs

loietermine the position in two dimensions, one has to add an extra LED. However, if two LEDs Kadded so that there is a pair of LEDs for each axis (horizontal and vertical), the discussion, and, Bled, the design of the system, are simplified. Four LEDs (a pair for each axis) are arranged so ibtheir irradiation patterns overlap (similar to Figure 4). If the LED array is rotated to the right, bout the vertical axis, the irradiation due to the LED on left of the array will increase and that on tetight will decrease. However, the irradiation due to the vertical pair will also be reduced, illout any movement in the vertical plane. However, the irradiation will reduce by an equal tmount; with appropriate processing at the receiver, no vertical movement will be detected.

USystem Configuration

ISbasic system configuration is shown in Figure 4. The user wears a Head-Mounted Unit (HMU) ta iommunicates with a Base Station. The HMU is responsible for sending signals either iKeraing its relative position to the Base Station or from which the base station can determine the ill's relative position and any mouse clicks. The Base Station is responsible for encoding clicks rdhead movement to the PC as standard mouse signals.

Toimplement a practical system, we can choose to have a single receiver and multiple LEDs, or Bltiple receivers and a single LED. We chose to have multiple LEDs and a single receiver kcause this solution is much cheaper. A second decision is whether the LEDs or the receiver aould be contained in the HMU. We chose to mount the LEDs in the HMU, as the alternative iRuires the received results to be relayed back to the Base Station by a separate LED link. This wuld require the overall system to have five LEDs and two receivers, whereas the option chosen lequires four LEDs and one receiver.

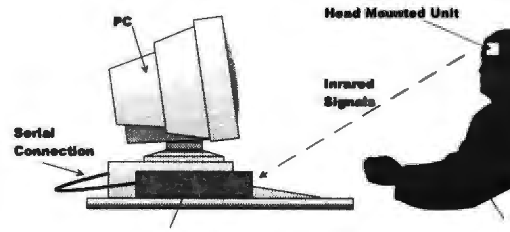


Figure 4: System Configuration

4.3 LED Mounting and Device Characteristics

We need to be able to align the LEDs so that their beams overlap to give ideal characteristics at the receiver. Consider Figure 5, which shows irradiation due to two LEDs (A and B), in the same plane against angular displacement at three different alignments together with a plot of the difference (A-B) against angular displacement. Figure 5(ii) shows the ideal arrangement, whose difference function is linear in the range between the two peaks. The separation of the two peaks determines the operating range of the device. If the spacing is narrow (Figure 5 (i)), discrimination is poor and, if the spacing is wide (Figure 5 (iii), the difference function become non-linear. It is difficult to analytically determine the best alignment. By experimentation at 'normal' operating distances a 15° mounting angle was found to be optimal for wide-angle devices (60° half-angle devices). It would seem that this angle needs to be precisely set for efficient operation. We discuss this issue further below.

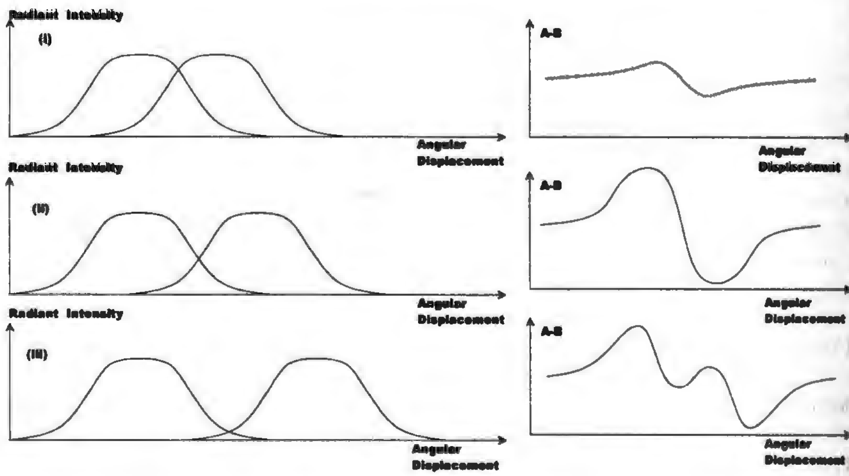


Figure 5: Alignment of two LEDs

The irradiation patterns shown in Figures 1, 2 and 5 are idealised; practical devices will deviate from this. Experience shows that their irradiation patterns may be asymmetric and they may also have patterns that have 'flat tops' or even irradiation patterns that have maxima that are not at the centre of the beam. The patterns often have very sharp cut-offs outside the centre of the beam. Moreover, even between devices of the same type there is considerable variation in irradiation pattern. These characteristics are thought to be due to the lenses of the LEDs, which have short focal length and may be subject to manufacturing variations. These problems compromise the accuracy of the measurement of angular displacement and may affect the optimal alignment for any given pair of LEDs.

These problems are not as serious as one might suppose. This is for two reasons. Firstly, we chose to operate primarily as a relative pointing device. This means that, in the simplest case, we only need to detect when head position has moved from the central deadband and the direction in which it has moved. Although we may chose to determine how far it is away form the deadband and control mouse speed so that greater deviations from the deadband cause greater speed. Secondly, and more significantly, there is a feedback loop. The user monitors the position of the cursor on the screen and moves his/her head accordingly. This makes the system very resilient to manufacturing variations in the LEDs and also to the alignment of the LEDs.

System Design

5.1 Transmitter Design

The transmitter consists of two parts, an array of LEDs that are mounted on the user's head (the HMU) and a small control box that can be clipped to the user's belt, rest on their lap or be held in their breast pocket. The HMU consists only of the LEDs and the control box contains the rest of the circuitry and its battery power supply.

The transmitter needs to send signals to the receiver, from which the receiver can determine the angular displacement of the head and also the state of the two mouse buttons. The latter are practically implemented as sockets to the Control Box into which users can connect switches according to their preferences.

A simplified schematic diagram for the initial design of the Transmitter is shown in Figure 6. The output of each of the LEDs and the signal received by the receiver are shown in Figure 7. Figure 7 shows each LED is modulated by a 16.384KHz signal at a 25% duty cycle. With appropriate receiver (see Section 5.2 below), the modulation removes interference from other light sources. The modulation frequency must be chosen so that it away from the line frequency of a computer monitor (generally 30-35KHz) and modern electronic ballast fluorescent lights (around 30KHz), otherwise interference may be picked up at the receiver.

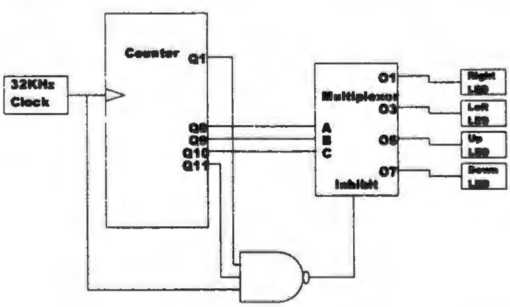


Figure 6: Simplified Transmitter Design

The transmitter uses a time division multiplexed protocol to transfer the necessary data (the four LED positions and two mouse button states) to the receiver, as shown in Figure 8. Firstly all LEDs are illuminated. The receiver uses this pulse as a synchronisation pulse. There then follows a gap after which the right LED and then the left LED are illuminated in turn. Next the state of the right mouse button is sent. If it is depressed, all LEDs are illuminated, as in the synchronisation pulse. The up and down LEDs are illuminated in turn, before the state of the left mouse button is sent. To ensure reliable synchronisation of the receiver, the transmitter waits for a period as long as the original symbol stream, restarting the transmission with the synchronisation pulse. This gives 16 transmissions per second and which is acceptable operation.

One important design consideration is the size of the synchronisation and mouse button state pulses. As these are the sum of all four LED outputs, they will, if full illumination of the LEDs is used, be considerably larger than the output for a single LED. As Automatic Gain Control (AGC) is used at the receiver and is controlled by the synchronisation pulse, this could mean that very low amplitudes are detected for the single LEDs. Thus the LEDs are not illuminated to their normal levels during synchronisation and mouse button transfers. The required level is difficult to calculate analytically and has been set empirically. Occasionally this means that the signal received for a single LED exceeds that of the synchronisation pulse. In this case the AGC simply adjusts and the system operates normally. The only requirement is that the synchronisation pulse is detected.

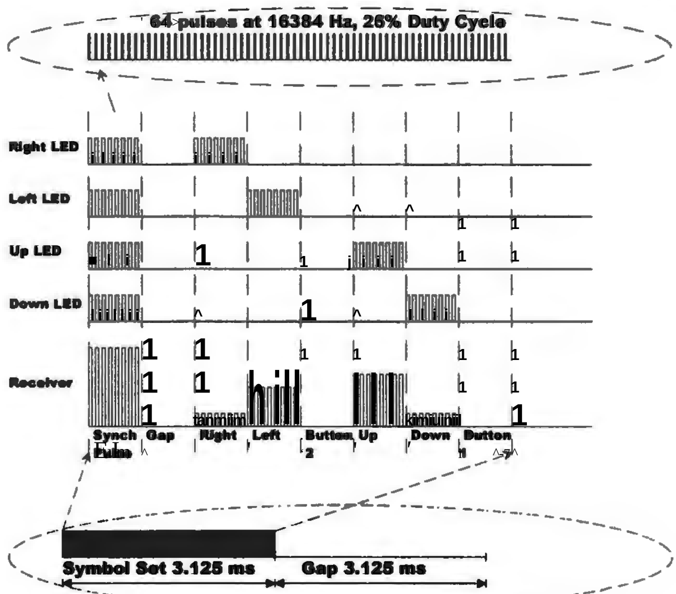


Figure 7: Transmission Protocol

The design of the final system is slightly different to that discussed above, although the principles remain the same. The transmitter is implemented using a PIC 16C54 microcontroller to control the illumination of the LEDs; this reduces system cost and power consumption compared to the discrete design implied by Figure 6. The modulation frequency is reduced to 8KHz, rather than the 16KHz discussed above. This means that the amplification stages in the receiver can operate more efficiently.

5.2 Receiver Design

The receiver (Base Station in Figure 4) is responsible for processing the signals from the HMU and encoding these as standard mouse control signals. A simple block diagram of the receiver is shown in Figure 8.

The light is received by a photodiode that is configured with a band pass filter. The band pass filter's design is intimately linked to the characteristics of the photodiode and it is present to reject extraneous infrared light. The filter chosen is a single LC resonator. This is low noise, has zero power consumption, provides a high impedance load for the photodiode giving a good signal at the modulated frequency, and gives excellent rejection of unwanted light.

The next stage, an HF preamplifier, has a maximum gain of around 1000. It is controlled using an AGC element. The AGC is generally controlled by the 'slow AGC' signal from the demodulator. This gives an AGC with a fast attack and a slow delay characteristic that minimises the changes in gain between sampling. However, if there are significant changes in the received signal, due, for example, to movement of the user toward or away from the receiver. There is a possibility that this AGC may not recover in time. Therefore a fast AGC is also provided. This is triggered by the microcontroller, in the gap between the synchronisation pulse and the other signals (see Figure 7), when the received synchronisation signal is outside a specified range. This ensures that the signals from the single LEDs are measured consistently.

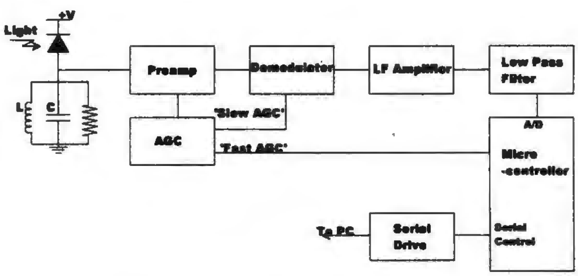


Figure 8: Block Diagram of the Receiver

A linear rectifier with a minimum voltage gain of approximately 3 then demodulates the signal. The signal is then amplified by a low frequency DC amplifier and passes through an active filter that removes the residual carrier component. This filter is designed to compensate for the slow rise time of the LC resonator and means that the target voltage is reached more quickly and thus sampling accuracy is much less critical. The signal is then fed into the microcontroller's A/D port.

The microcontroller is responsible for sampling the incoming signal, determining head position and switch status and sending appropriate mouse control signals through its serial port to the PC. We use a PIC 16C71 microcontroller. This has a more than adequate performance, a low parts cost and low power requirements. The analogue processing is carried out using a quad CMOS opamp (TLC274). This is also low cost and more than adequate for work at 8KHz.

5.3 Microcontroller Software

This section focuses on the processing required to determine the mouse movement. As noted in Section 4.1 the determination of angular position in one dimension, say the horizontal plane, is given by A-B (where A is signal derived from the left LED and B the right). However, this is not ideal for a number of reasons.

Firstly, as noted in Section 3, the signal derived from the left and right LEDs is also dependent on the movement in the vertical plane. However, one can easily scale the difference (A-B) by dividing by A+B. In practice, however, this approach has problems, because, as A-B/A+B becomes small (toward the margins of operation) noise and quantization errors start to dominate. We therefore need to multiply the expression A-B/A+B by some factor. Experimentation shows that, if one uses a factor of A+B, i.e. uses only the difference function A-B, the system works acceptably well provided that the result is processed in the correct way, as discussed below.

Secondly, we need to consider the operational characteristics of the device as perceived by the user. It is desirable that, as the head moves away from the deadband position, the cursor should move more quickly. Thus, close to the deadband the user has fine control, further away from the deadband, the cursor should move much faster. This requires a non-linear response. An associated problem is that we require a deadband. If we directly use the result of A-B, head positioning becomes critical and there is significant 'cursor creep'. These problems can be overcome if the result of A-B is used as an index into a look up table, which effectively encodes the rules for the deadband and mouse speed. Indeed, the system contains a number of rule sets that govern the size of the deadband and the speed of the mouse. These can be selected by DIP switches in the Base Station.

Whilst this arrangement may not seem ideal, in that movement in the vertical plane is not factored out of the determination of horizontal position, and *vice versa*, the system works very well. This is another example of the user's ability to use visual feedback of the cursor position to correct for defects in the design of the system and its components (see Section 4.3). The user does not realise that he/she is compensating.

6. Performance

The system has an effective transmitter to receiver range of 15cm to 10m and is tolerant of changes in transmitter to receiver distance when the system is in use. If an array of three photodiodes is used at the receiver, the effective angle of use is better than 120° in the horizontal plane. This gives considerable freedom of seating position and movement.

The transmitter/HMU requires 2mA at 3V. For two alkaline AA batteries this gives a working life of 1300 hours, which is around 160 eight-hour days. The Base Station is powered directly from the PC's serial port and requires no other power source. The system will work with laptops, but obviously puts a small additional load on their batteries.

It was noted earlier that the device operates as a relative pointing device. However, for certain applications an absolute pointing device is useful, one in which the cursor follows head position. For the reasons given earlier, it is not possible to construct a device of this sort with a high degree of positional accuracy. However, our device can be used as an absolute pointing device with a resolution of 256 x 256. In this case the system does not emulate a Microsoft compatible mouse and consequently a custom mouse driver must be loaded into the PC.

7. Status

At the time of writing, our head operated joystick exists as a pre-production prototype. It has been used with around 30 users, a small number of whom have been disabled. It is expected that the device will be commercially available by the time that these proceedings are published.

8. Acknowledgements

The authors would like to thank the Southern Trust and the Guide Dogs for the Blind Association for the funding of this work. We would also like to thank the Foundation for the Communication for the Disabled for assistance in the evaluation of the device.

9. References

[1] RIGOL, M. V., Determining Head Position for Mouse Emulation, MSc Dissertation, UMIST, 1996

[2] HeadMouse Head-Controlled Pointing for People with Disabilities, Origin Instruments, Texas, USA

[3] COLINEAU, J *et al*, Laser Mouse In: I Placencia Porrero, R. Puig del la Bellacase (eds) The European Context for Assistive Technology – Proceedings of the 2nd TIDE Congress. IOS Press Amsterdam, pp 290-293, 1995

[4] GIPS, J., DIMATTIA, P., CIURRAN, F., X., OLIVERI, P., Using Eagle Eyes - An Electrodes Based Device for Controlling the Computer With Your Eyes - To Help People with Special Needs, Proc. 5th International Conference on Interdisciplinary Aspects on Computers Helping People with Special Needs, pp 77-83, 1996

[5] EVANS, D. G., PETTITT, S., BLENKHORN, P., 'A Head Operated Joystick' Proc. 5th International Conference on Interdisciplinary Aspects on Computers Helping People with Special Needs, pp 84-91, 1996

[6] R ULIG, Making eye contact as everything clicks, Daily Telegraph Connected (15 July) 1997

[7] NIELSEN, T., Opto-Electronics by Design, Electronics and Wireless World, pp 364-360, May 1994

Session IX

Tele-learning & Teleworking

u

TELE-LEARNING: OVERCOMING DISTANCES AND DISABILITIES

Norman Coombs

Chair of EASI (Equal Access to Software and Information)
Professor Emeritus Rochester Institute of Technology

Richard Banks

EASI Electronic Resource Manager

Tele-learning or distance learning is, probably, the hottest topic in education today in the United States and Canada. It has captured the attention of universities and businesses alike as a means of delivering education and training that overcomes the barriers of time and distance. EASI (Equal Access to Software and Information) believes that distance learning technologies also provide unique opportunities to deliver such education and training to students with disabilities. However, if such systems are not designed with this population in mind, distance learning technologies may erect new and needless barriers to participation by students with disabilities.

Definition:

First, distance learning is a misleading label. What began as a set of technologies to deliver education to students who were located some distance from the college or university became used as well by students living near or even on the campus. While some users have tried to change to the term, technology-mediated learning, the term in popular use in the United States today is "distributed learning".

Second, most of these systems are either digital or use telecommunications to transmit content. They include telephone conferencing, video conferencing, computer conferencing, e-mail, fax and even radio in some remote rural districts. One way to lock at these different systems is to divide them into those, which are synchronous, and those which are asynchronous. The more synchronous systems like interactive video, computer chat rooms and telephone conferencing usually try to simulate the classroom. Asynchronous systems like computer conferencing and e-mail try to adapt the content presentation to the strengths and weaknesses of the particular system.

Advantages and disadvantages of Distance Learning

To some extent, the strengths and weaknesses of distance learning differ according to which specific technology is being considered. The subject matter is another ingredient in how well distance learning works. However, the most significant factor in measuring the effectiveness of distance learning is the teacher. This is no surprise as it is the same factor that is important in evaluating the effectiveneii of face-to-face teaching. To the surprise of many, distance learning has generally been found to increase interactiveness between students and teacher. It also has deepened the sharing of ideas and made many students more will to admit and share their problems. This short presentation does not permit time or space to expand on this point. These interactive factors are most important in subject areas that lend themselves to discussing ideas and theories and are less significant in topics that are largely based in rote memorization and which are fact-based.

Frequently, these distance systems provide a mix of guided learning and independent learning. This permits each student to work at his or her own pace. These systems are adaptable to a variety of learning styles. This individualization also means that these systems are usually adaptable to the learning needs of students with a variety of disabilities.

Distance Learning and Students With Disabilities

The computer which is integral to most of these distance learning technologies, as we all know at this conference, can be adapted to include and empower students with disabilities in today's information age. This also means that it has the potential to include these students in distance learning. Even more when we remind ourselves that this is much more than distance learning, it has the potential to provide better access to modern, technological education than ever before. We also all know that improper design may create roadblocks to learning that could become a serious problem. Many institutions are aware of the benefits and problems and are designing their systems with this in mind. Others, in the rush to get on the distance learning bandwagon are not taking these needs into consideration.

Western Governors' University

The governors of 17 of the American western states, headed by the Governor of Colorado, agreed to jointly develop a distance learning university, and more states are expected to join. Actually, it is not so much a new university as it is a consortium of existing programs given a new impetus and a broader mission. During its development WGU created a focus group to study the need to provide access to its programs by students with disabilities. The report stated that WGU would want to provide quality access for disabled students because modern technology made this possible, because US civil rights law required it and, because as public institutions, they would want to provide services to all their citizens.

The recommendations of this report are divided into the following eight topics:

1. Information technology systems which primarily means computer systems and computerized information which need to be designed to be accessible by adaptive computer software and hardware-screen readers, screen magnifiers, voice recognition, alternative keyboards and onscreen keybords.

2. Regional and local centers which means that WGU will have some regional, physical locations where students can come for help and to access some of the needed technology, and these sites and their equipment will need to be fitted with a spectrum of adaptive technologies.

3. University publications will have to contain a statement outlining WGU's commitment to access, and these publications will need to be available in alternative formats.

4. "Academic courses, course materials and delivery technologies will have to be made as accessible as possible. Some of this requirement will be accomplished by point 1 which discusses access to information technology. Even there, however, individual professors will frequently design their own web pages and other interfaces and will be unaware of the special needs of disabled students. Frequently, well designed course presentations can adequately provide material in a format so that it is understandable by students with disabilities and minimize the need for special adaptations.

5. On-line library services are becoming of universities even when there are no distance learning courses. Besides the need to provide interfaces that are accessible as described already, the library must take care to store information in generic, display-independent formats that can be used by students with disabilities. Text stored as pictures render those texts useless for someone using a screen reader. Coincidentally, they will not be searchable either.

6. Student services must be provided for distance students as well as for those living on campus. Many of these delivery systems, designed with care, will also provide a better level of service for those with disabilities.

7. Training for disabled students on the use of both the distance technology and of their adaptive technology is important. Unless these technologies are so common in previous education that students can be expected to be familiar with them, then training on the technology is important. Without it, students struggle with the technology, and this only increases the difficulty of learning the academic material.

8. A coordinator of services to distance students with disabilities is an important position that must be created. Someone needs to be responsible for overseeing the integration of these complex systems. If everyone is responsible, then, in fact, no one would be responsible.

The Western Governors' University is in the early stages of being developed. Only time will tell how well it puts these principles into action. WGU is only one of the exploding number of distance learning institutions. Professionals with an awareness in this field need to alert other universities of the opportunities and challenges.

Other Distance Learning:

Many American businesses, especially those which are multinational, are making increasing use of distance learning to provide training for their employers. Other smaller organizations that are trying to impact learners at distances also use this technology.

EASI (Equal Access to Software and Information) has three online workshops which have reached participants in over 30 countries on six continents. See information on the workshops at http://www.rit.edu/~Easi and select the online workshop link. Adapt-it is a four-week workshops designed to provide institutions with the information they need to make their computer and information systems available to people with disabilities. EASI-SEM is a similar workshop with a focus on helping to mainstream students with disabilities in the fields of science and math. It has recently expanded to include information relevant to special education teachers of students in grades k-12-pre college). EASI-WEB is a four-week workshop on web design. It explains the needs of disabled web surfers. It explains features in HTML code to help provide access for them. It also deals with making multimedia accessible by providing both captioning and video descriptions. All workshops make extensive use of the web because it is a universal technology, and it readily promotes a lot of interaction between participants and instructors.

EASI has recently obtained the ability to do streaming audio and video.
These technologies can either shut out the blind and deaf or, if audio and video are used to provide redundant information, they can be used to increase access for these groups. Check the EASI web http://www.rit.edu/~easi and especially the link, "EASI Listening" which has some multimedia examples. EASI is beginning to add multimedia slide shows to its online workshops in which it integrates graphics, text and audio in ways intended to increase access for all.

Bio for Dr. Norman Coombs

Coombs is a professor of history and has been involved in distance learning for more than a decade. Being a blind professional, he makes extensive use of adaptive technology which led to his becoming the Chair of EASI (Equal Access to Software and Information) and directing a National Science Foundation Grant for EASI, an affiliate of AAHE. Coombs is also providing consultation on disability access issues for the Western Governors' University. He has presented on both the topics of disability access and of distance learning across the US and in Canada, England, Switzerland, Hungary, Turkey.Coombs co-authored with Carmela Cunningham INFORMATION ACCESS AND ADAPTIVE TECHNOLOGY, Oryx Pres, 1997.

Bio for Richard Banks

Banks has been an adaptive technology consultant for University of Wisconsin-Stout and has served as a consultant on adaptive technology' in Thailand for Ratchasuda College - Mahidol University Bangkok. He is EASI's electronic resource manager and responsible for its extensive web pages and responsible for overseeing several of EASI's listserv discussion lists. Banks also assists Coombs in teaching an online workshop on adaptive computing which has reached over 500 people in more than 24 countries.

Teleworking for Disabled People: Pitfalls and Ongoing Challenges

Michael Pieper, Henrike Gappa and Stefanie Mermet

Abstract: The TEDIS-project (Teleworking for Disabled People) is a publicly funded research & development effort of the research group on Human-computer Interaction of the GMD - German National Research Center for Information Technology. The project was targeted at implementing a suitable teleworking environment for disabled end-users thereby taking into account usability issues as well as a social technology assessment. The promising potentialties of adaptive and adaptable user interfaces for the vocational integration of the disabled are exemplified as well as the disappointing truth that their participation in todays working world with long-termed professional occupation is almost impossible to achieve. However, there are indications, that concepts of a center-based organization of the labourforceof disabled people could resolve this antogonism.

1. Introduction

Disability is mostly to a different degree restricting the mobility of handicapped fellow citizens. In Germany 33% of all legally acknowleged severely handicapped people suffer from a functional impairment in the upper and/or lower limbs, which often causes a - so to say - primary mobility restriction. Secondary mobility restrictions have also to be taken into account, which result mainly from receptive handicaps restricting spatial orientation, e.g. blindness or vision-impairment. On that background, TEDIS concentrated on telecooperation technology to overcome the often restricted mobility of disabled persons, which is particularly detrimental to their professional integration.

The TEDIS-project was carried out in cooperation with the internationally renowned FTB (i.e. Research and Development Institute for Assistive Technologies). The FTB contracted two severely physically handicapped end-users serving as pilot users within the TEDIS case study. They worked for the marketing department of the FTB with support of a telecooperation environment being developed at GMD's Institute for Applied Information Technology. In the course of the case study the two teleworkers created a catalogue, consisting of mixed document pages with embedded text, graphics and pictures illustrating devices of technical special needs equipment, provided by the FTB.

2. Technological special needs adaptation of a teleworking environment

Technologically, TEDIS offers technical integration of standard and special needs equipment, closing technological gaps between standard and special needs equipment by own technical developments. The TEDIS teleworking environment consists of the internet-based telecooperation system BSCW (Basic Support for Cooperative Work) with adaptable GermanBig user interface, the Speech input system Dragon Dictate 2.2, a head pointing device („Headmouse"), SmartCard technology (contactless chipcard) and a WWW browser with special needs adaptation.

The internet-based telecooperation system BSCW has been developed by GMD's research group on Computer Supported Cooperative Work (CSCW) [1]. The system supports collaborative work over the Internet. One of the enhancements derived from a first usability study of the BSCW-system was the implementation of the additional feature „GermanBig". When the user selects this feature all functional icons are enlarged. This facilitated navigation by mouse significantly for end-users with fine-motor impairments. It is not only easier for them to locate the cursor on top of a button, perceptual problems with icon clusters can also be prevented. Nevertheless it also turned out that there is a trade-off between facilitating navigation by icon enlargement and a reduced overview of the workspace [2].

For SmartCard technology plastics cards as big as a credit cards are used containing a chip to store personalized user-profile data. The application domains for such a personalized technology can be ubiquitous (PCs, kiosks). TEDIS implemented a prototype of this technology into its teleworking environment to enable end-users to configure default options and preferences corresponding to their special needs independently since they can be set automatically.

Additionally TEDIS developed a prototype WWW browser called WebAdapter, which provides new special needs adaptations for physically handicapped, blind and visually impaired end users up to now not being offered by standard web browsers. This browser is more comprehensively circumscribed in another chapter of these conference proceedings [3].

3. Pitfalls: Examples from the TEDIS case study

In the course of the TEDIS case study focal points of the accompanying technology assessment related to social issues of telework, the task organisation, the work flow, the task related communication, and the already mentioned usability of the telecommunication environment. At this point it should be underlined that the results reported from the TEDIS case study are not seen as representative. It was more the goal of TEDIS to gain qualitative rather than empirical data.

Generally the two disabled teleworkers had a clear preference for home-based telework mainly because it gave them more flexibility in determinig their working hours individually. In turn the flexibility of individually determinable working-hours demands for a high degree of self-organization not to be underestimated. Telework in general must be managed much more by reaching definded objectives rather than allocating and controlling the time devoted to reach certain objectives.

In the beginning of the case study, teleworking was not as efficient as it turned out to be. Partially that was due to technical restrictions in speed and reliability of data transfer which could be resolved later on. Thus, initially a visitor from the FTB acted in temporary manner as an intermediator for organizational task-oriented *face-to-face communication* between the employer and the two teleworkers instead of using telecommunication facilities. It was mainly the investigation of the organization of the teleworking process that revealed some obvious deficiencies in workflow, time-management and task-related communication. Overall, not deficient telecommunication - or deficiently adapted standard and special needs technology, but a suitable concept of organizational procedures for a mid- to long-termed *management by objectives* in a socio-technological telecooperation environment has been proven crucial for the success of teleworking.

4. Challenges: The gap between TEDIS intention and reality

Social issues of home-based telework refer mainly to the fact, that support of disabled people by telecooperation technology might be detrimental to their social integration and isolate these people in a world of technical devices which they are able to handle, but which secludes them from social life. First of all, in our analysis of the information gathered by guided interviews, closed and standardized questionnaires [4, 5], no evidence could be found for increased occurences of feelings of isolation. In the beginning of the case study it became apparent that the teleworkers most of all put emphasis upon a challenging type of work. Least of all, they used to be eager on forming social contacts through work. Concerning the latter issues, it is to mention that in the course of a final interview of the TEDIS-project, both teleworkers showed differing attitudes. Both are still definitely in favour of working as a teleworker, however, they would now prefer meeting with other teleworkers at least once a week. They would like to form contact with other people belonging to the same professional subgroup of teleworkers they seemed to have identified with.

Resuming our social technology assessment has shown, that telework can be a feasible form of work organization for people with a severe mobility-impairment. However, in the mid-term future it is predicted that severely handicapped people, who acount for 6.5 million people in Germany, will be particularly affected by a still rising or at least high-level unemployment rate [6]. The current unemployment rate amongst the severely handicapped in Germany accounts for more than 16%, compared with an overall average of about 11%. Surveys [7] show, that the majority of this group is very motivated to get work. They are even open to change their profession regardless of originally gained qualifications rather than staying unemployed.

Examining statistical data about the nature of most handicaps, it becomes obvious, that 33% of all severely handicapped people in Germany, suffer somehow from a functional impairment reducing their motor capabilities, which often causes restricted mobility. Arranging worksites as teleworking environments should offer an opportunity for compensation and competition on a tense labour-market for this target group. However, despite a lot of efforts and projects to (re-)integrate mobility-restricted fellow-citizens into the societal worforce, we only know of very few examples which have been successful in the sense that telework has founded a long-termed professional occupation. If need be, telework has created *work* but not stable *workplaces*, secured by labour legislation concerning social insurance, pension funds and the like. Industrial sociologists agree, that especially in the services sector, telework will become an important new part of the occupational culture of modern societies, but the teleworker will mainly be self-employed, or - if the worst comes to the worst - pretendedly self-employed, which means that the teleworker highly depends on only about one or one customer.

It is thus the probability of lost customers, which not only disabled people being involved in teleworking projects aiming at their vocational (re-)integration will have to take into account: if not having been self-employed, the disabled teleworker may not be able to make claims out of a working contract, because there is none. Generally two reasons may contribute to this. Either the terms of contract for getting public funding for a teleworking project do not allow spending part of the funds for contracting the disabled teleworker, because funding is for research and development only and not for vocational (re-)integretion, or there is no contract, because otherwise the wages to be paid out of it are legally demanded to be balanced with payments out of welfare funds, most of the severely disabled are entitled to.

For reasons of pure national economics, this should by no means be criticized at all. Real life experiences of the disabled however withdraw the ins and outs of this matter from pure national economics reasonning. Especially the statistically proven experience of telework creating work but no stable workplaces combined with the mostly nerve-racking struggle to apply for legally adequate and most of all timely payments out of welfare-funds diminishes the motivation to skip monetary support in favour of an indeed paid but presumably only short-termed employment.

5. Conclusion: Center-based organization of the labourforce of disabled people

The social organization of disability-related telework has thus to acknowledge a disappointing fact. It has to deal with the antagonism of people with severe handicaps and mobility restrictions who would very much like to participate in the working world of our society and the economic situation of aggravating competition on the labour market being largely settled at the expense of personal socio-economic security and partial demolition of welfare systems, disabled citizens [illegible] special extent.

Teleworkcenters, at best attached to conventional rehabilitation centers, may be the answer. They can best come up with a sufficient number of teleworkers, whose combined labourforce builds up a critical mass for reciprocally compensating social and economic security expenses of each other.

Furthermore homebased telework environments as shown in TEDIS, or even neighbourhood satellite offices shall also be attached to teleworkcenters for disabled people who are unable to attend somehow. That is why we actually ascribe the term „center-based" instead of „centered" to our approach, because telework from „outside" needs center-based care. Or, as both our pilot teleworkers claimed it: Homebased telework in any case needs to be centrally supported by an all day accessible mobile technical service and a full time central online management on a regular base with temporary organizational face-to-face meetings.

References:

CARRUTHERS, S., Anne HUMPHREYS and Jim SANDHU, Out of the net. Teleworking and people with special needs, A report prepared for the European Commission's TIDE 166 INDICES Project, Special Needs Research Unit (SNRU), University of Northumbria, Coach Lane Campus, Newcastle upon Tyne 1992

DOERING, N., Isolation und Einsamkeit bei Netzbenutzern?", Oeffentliche Diskussion und empirische Daten, Stuttgart 1995 (ftp;//ftp.uni-stuttgart.de/pub/doc/networks/misc/ netz_und_einsamkeit)

MUMMENDEY, H.D., „Die Fragebogen-Methode". Göttingen 1987

GAPPA, H., R. OPPERMANN and M. PIEPER, Certtifying Web Accessibility for the Handicapped by ISO 9241 Conformance Testing, in proceedings of the 3rd ERCIM Workshop on User Interfaces for All, Obernai 1997

HENNIGES, H. von, IAB Projektbericht, Arbeitsmarksituation und Merkmale von arbeitslosen Schwerbehinderten, Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit, Januar 1997

[3] HERMSDORF, D., WebAdapter: A prototype of a WWW browser with new special needs adaptations, draft version of a paper submitted to ICCHP '98, Sankt Augustin 1998

http://bscw.gmd.de/

Opportunities for disabled jobseekers with

"Online Teletraining"

Mag. Dr. Georg Tschare

1.Idea and programme

The project "*Online teletraining*" started in March 1997.
The target of the project solicitors was and is to integrate disabled people by the means of an intensive qualification programme in combination with appropriate efforts directed at finding long-term employment.
The primary target group consists of people whose entry into the labour market is hindered or prevented by a disability restricting their mobility and for people who hold a confirmation of a disability level of at least 50 percent issued by the Federal Social Affairs Office.
The qualification programme lasts about four months and focuses on training computer skills, which will put the course participants into the position to do a job also in the form of telework.
Especially for people with mobility-disabilities telework represents a singular chance for integration into the primary labour market.
It seems only right that the course is done by teletraining, i.e. distance learning.
Firstly, this form of training takes the participants mobility problems into consideration, and secondly it represents the best preparation for future teleworking specialists.

2.Contents and mode:

The main emphasis of teletraining lies on developing computer skills, primarily with standard applications of office software.
Training in data transfer, telecommunication techniques and telematic in general constitute a focal point of the entire course.
Participants shall also receive an introduction into graphics programmes.

...net application such as the use of various online-services, e-mail as a means of communication, ...ML-programming, etc. have become an additional focal point.

...nsive telephone-training is a condition to be successful as a teleworker.

...ther important step towards integration, besides the technical skills, is definitely the support ... finding jobs. This support comprises application training, advice, the organisation of ...rmation seminars for interested companies, etc.

...s therefore not only computer skills what we are trying to convey.

...training on a one-to-one-basis enables us to create an individual, tailor-made programme for ...ry single participant. This is how we can suit the capabilities of every trainee and also his/her ...ds as regards the prospective jobs. If a company is interested in a particular trainee additional ...lifications for the future job can be offered at any time.

3. Chances

The *"Online method"* and especially its tailor-made and individual training concept is the result of a ...rket research on potential employers.

The biggest chances for jobs (vacancies and new jobs) lie with small and midscale enterprises. It is ...m who hope for an additional advantage in competition and for new chances on the market by using the Internet. For that they need well-trained employees who can also provide the company with know-how on the new media possibilities.

...working offers new chances on the free labour market especially for those people whose ...bility has so far hindered them from a regular occupation. A unique chance for people with ...stricted mobility but good qualifications.

...ic conditions

The growing importance of telework in the modern world of work is undeniable and yet has only ...ntly been established. Telework still possesses an aura of modernity and attractivity due to its widely expected potential of cost reduction. It is therefore an interesting subject for businesses decisions on implementations of this type of work.

"Online teletraining" offers skills to trainees which are advantageous for their competition against ...n-handicapped jobseekers.

At the same time one has to be careful not to paint wrong, not fullfillable and too optimistic ...pectations, which are still widely shared.

These unfullfillable expectations are both on the side of the employees and the employers – extensive information has to be provided.

Hereby, the aim of *"Online teletraining"* is clearly defined:

To show new job possibilities for disabled people which have emerged from the huge expansion of information- and communication technologies.

The employment of disabled jobseekers meets remarkable resistance from many companies. The Reasons named are legal conditions, such as the extended protection against dismissal which are meant to provide a better job security for handicapped people. Furthermore, one has to fight prejudices, such as more frequent sick-leaves.

An important means to convince enterprises to offer jobs to handicapped people is to inform them on the subsidy by the government and local authorities as part of the active labour market policy.

4.Aim

The project "*Online teletraining*" is an attempt to significantly improve the chances for handicapped people, often unemployed for long time spans, by the means of investing in the qualifications with new and innovative areas on the job market.
This attempt to integrate disabled people can be successful if it is accompanied by all the other measures mentioned above.
The results prove the necessity of such a project which can be seen as a mission for the future.

HUBER, H., KRASSNIG, E., TSCHARE, G., Projektstudie Teleschulung, Wien 1996.

A TELE-LIBRARY SERVICE FOR THE MOTOR IMPAIRED: A QUALITATIVE STUDY ON THE MAN/MACHINE INTERFACE.

Isabel Van Mele, Kirstie Edwards, Mieke Verheust,
Arthur Spaepen

Abstract

On the road to UMTS, a need for ergonomic user centred design is emerging, especially for services on portable devices where screen space is restricted. This paper describes the development of a multimedia tele-library service accessible over a mobile network for people who are motor impaired, using iterative user trials to determine the interface design. This prototype allows consultation of a catalogue service and ordering and discussion of books remotely with library personnel through a real time video link. Initial user trials stressed the need to identify optimal on-screen sizes of interface objects, which are practically usable for people who are motor impaired where screen space is limited. Recommendations from these studies have been applied in the prototype's user interface, which will be demonstrated. Future research will evaluate the tele-library prototype implemented in the homes of students who are motor impaired, with mobile access from a wireless LAN.

1. Introduction

Mobile Telecommunications offer great potential for people with special needs in communications, education, employment, entertainment and everyday living. Currently, however, there is a shortage of appropriate terminals, services and intelligent networks, so that useful UMTS applications for these users cannot be easily realised. Until useful applications are available, the telecommunications industries will not provide the necessary infrastructure. The UMPTIDUMPTI project aims to demonstrate useful UMTS applications for users with special needs. In this paper, a multimedia tele-library service for people who are motor impaired will be discussed.

2. A multimedia tele-library service: rationale

Wheelchair-bound students have difficulty in accessing buildings such as libraries, so that being able to look up and order digital copies *electronically would save considerable time and effort.* This will only be possible if the users have input devices suited to their abilities, and providing the service has an accessible user interface, and is truly available from anywhere at any time. A portable terminal with a small camera attached will be used, and a wireless LAN infrastructure to provide mobile coverage for service provision.

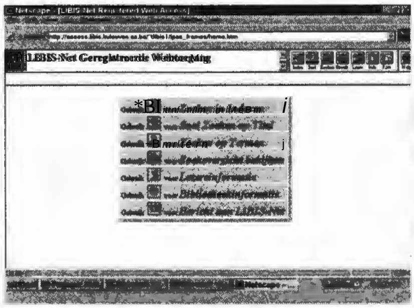
3. Functional Specifications

The service provides multimedia information across a mobile network to users who are motor impaired, and comprises: the search for information in the digital library catalogue, retrieval of information over a mobile network, and a real time ordering facility at the library help desk, [2,3,4].

We used LIBIS-Net digital library catalogue [6], which has two access options available on the WWW. Users will be able to request viewing and scanning of documents, through real time video-conferencing to a manned help desk. Scanned pages will be sent by e-mail to the user, and also saved on a server database for future use. The Advanced Technology group of White Pine Software are cooperating in the integration of video-conferencing in the service prototype. The web browser, the file transfer and the video-conferencing function will be integrated into a single stand-alone application through developments in Java.

4 User Interface: LIBIS-Net

Two user interfaces of the LIBIS catalogue available on the Internet were tested [5], to determine which design was preferable to users who were motor impaired, and to identify any problems.



Graphic 1 : Graphic user interface of the commercially available LIBIS catalogue

Standard deviations increased with the difficulty of the task, emphasising the magnification of difficulty for users who are motor impaired. Recommended improvements to design address *reducing the motor actions* necessary by improving *consistency* in design, *optimising instinctive awareness* of how to use the system, providing *informative feedback* and *instructive messages*.

To minimise screen space used for the toolbars of existing browsers, a browser with only the basic functionality necessary to use the library database has been developed in Java. HTML Forms have been adapted in-line with the above findings specifically for the library and stored on the Web server. They contain the substitution codes for the CGI scripts, and will subsequently become the user requested HTML documents.

Graphic 1 : Graphic user interface of the multimedia tele-library service

This research highlighted the need for guidelines on the minimal size of on-screen objects usable by motor impaired users. Small buttons are an advantage where screen space is limited, but there is inevitably a threshold below which a button's size becomes too small to use practically.

5. User Interface: On Screen Objects

Second trials [1] were designed to assess this minimum size of button and scrollbars/scrollbox practical for both able bodied and motor impaired users. Decreasing the button size from 225 mm^2 to 9 mm^2 has little or no effect on use by people who are able bodied, but affects the selection time and accuracy of people who are motor impaired. Reducing the button size to under 75 mm^2 affected performance by subjects who were motor impaired in these trials. All screen objects in the tele-library service will be adapted accordingly.

6. User Interface: Videoconferencing

In these trials subjects who are motor impaired made videophone calls with text transmission. User feedback and researcher observations identified problems people who are motor impaired experience with the software. Usability appeared to be related to computer literacy of the subjects rather than level of motor impairment. Recommendations for improvement included larger buttons, scrollboxes, and mark boxes, preferably over 81mm^2, minimising the distances which need to be moved with the cursor, and avoiding intricate tasks such as using drag bars. Optimising instinctive awareness of how to use the system, providing informative feedback and prompting messages on how to progress, and automatic procedures where possible would also eliminate costly effort in terms of motor actions for these users. The user interface elements of the videoconferencing facility in the tele-library service will be adapted accordingly.

7. Acknowledgments

This research is part of the project, Using Mobile Personal Telecommunications Innovation for the Disabled in UMTS Pervasive Integration, sponsored by the European Commission (AC027), under the Advanced Communications Technology and Services Programme. Parts of this research have been joint initiatives with LIBIS and the Advanced Technology Group of White Pine Software, whom we thank. We are also very grateful to all the subjects in these trials.

8. References

[1] AC027 E.U. DGXIII, "Determination of optimum button size for people who are motor impaired.", Public Deliverable 28, Pages 20-24, 1997

[2] AC027 E.U. DGXIII, "Functional Description of Proposed Retrieval Service for Motor Impaired People", Public Deliverable 10, Pages 22-26,1997

[3] AC027 E.U. DGXIII, "Technical description of proposed service for motor impaired people; multimedia information retrieval service", Public Deliverable 13, Pages 14-20,1997

[4] AC027 E.U. DGXIII, "Developments For Motor Impaired People - KUL", Private Deliverable 15, Pages 40-46, 1998

[5] EDWARDS, K., VAN MELE, I., VERHEUST, M., SPAEPEN, A., Evaluation of user interface design to optimise access to a library service over the internet for people who are motor impaired, Information Technology and Libraries, 1997.

[6] REGENT, A., "LIBIS-Net and CALIBRE", TOIDC - Belgacom, Brussels, 1995.

[7] VAN MELE, L, EDWARDS, K. ,VERHEUST, M., SPAEPEN, A., Mobile access to WWW services for university students who are motor impaired : a qualitative study on the man/machine interface, Proceedings of the 16th International Symposium on Human Factors in Telecommunications, Oslo 1997.

9. Biography

Isabel Van Mele has a B.Sc. in Mathematical Physics, and an M.Sc. in Biomedical & Clinical Engineering. Since 1994 she has been working on the project, "Using Mobile Personal Telecommunications Innovation for the Disabled in UMTS Pervasive Integration, which is sponsored by the European Commission (AC027), under the Advanced Communications Technology and Services (ACTS) Programme.

Human-Computer Interaction (3)

A VOICE DRIVEN SYSTEM TO SUPPORT DISABLED PEOPLE FOR USING COMPUTERS

Elaheh Pourabbas, Maurizio Rafanelli[1]

Abstract

In this paper the authors discuss a "free hand access system" on computer by means of which it is possible to use different types of software applications using, simply, a vocal interface supported by a particular software. The design and implemetantion of this system are part of the Italian Project on Telemedicine which is addressed to support the disabled people in computer using activities. The project uses also Telematic services for teleconferences and carries out a set of operations for disabled people.

1. Introduction

Disabled people were often considered as a minority and very often margined group, particularly with regard to the technological aspects, such as, for example, the use of computers. If, from a given point of view the technology of the telecommunications obtained important progresses in many fields, from another point of view, only a very few people are considered with making these technologies accessible to disabled people [1].

At present the Telematic networks are able to connect millions of computers, which communicate among them, so that we can consider, in every aspect, such computers as "means of Telematic communication".

The technological evolution, and in particular, the above mentioned communication means, facilitated the social relations and the work of people without a physical handicap (or "physically normal people"), but, paradoxically disabled people have become even more isolated and elderly people have greater problems with ageing [2, 3, 4].
Actually, it is impossible for a disabled person to be part of a working context. The results of this fact are both his economic insufficiency and consequently the increase of the costs for the community, and psychological isolation.

[1] Istituto di Analisi dei Sistemi ed Informatica-CNR, Viale Manzoni 30, I-00185 Rome, Italy

In the last few years an effort aiming to change attitude of indifference towards the socially weaker population needs has been made. In fact, disabled and elderly people, far from being a minority group, represent a significant percentage of the population, which will increase as time goes on more and more.

Moreover, it was demonstrated that a large part of them can recover to a normal activity, simply by using common devices which the advanced technology is able to provide.
Recent research projects [for example, TIDE (Technology Initiative for Disabled and Elderly people), of the European Communities] estimated that at present in Europe there are at least 80 million people who have difficulties in working or in their daily life, and to a handicap and/or to aging.

These people wish strongly to keep their autonomy, and this desire creates a strong need for tools and services able to provide them with the assistance for maintaining their daily activities.
In these last years, different researchers studied and proposed tools for improving the life of disabled people. Many areas were investigated and, for each type of disability, hardware and software tools were proposed [5, 6, 7]. In this paper, the handicap of the inability of hands, which refers to the proposal discussed in [8], is studied, a user-friendly interface based on the vocal communication (instead of using the keyboard) is discussed and a prototype of a system able to support different functions is illustrated.

This research is supported by the Italian National Project "Telemedicine" and it is one of the three goals "Help communication systems for the disabled" [9, 10].
The paper is structured in the following way: In section 2, the system architecture is illustrated. In section 3, the temporal phases of the project are illustrated. In section 4, the principal characteristics of system, which lead to the description and the illustration of its functionality, are discussed. Section 5 concludes.

2. The System Architecture

The general aim consists of the definition and the implementation of a system for the access to peripheral units, telecommunication and Telematic "free hands" services (that is, without the use of the hand) by vocal commands, specific for disabled people. From the hardware point of view, a phonetic terminal is implemented, which is characterised with the possibility to be used through vocal commands. The hardware configuration of the system is shown in Figure 1.

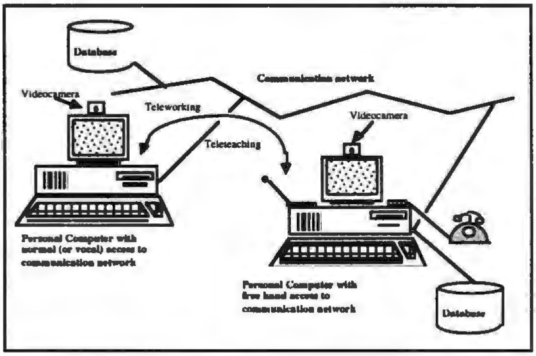


Figure 1. The system architecture

The system consists of some prototype stations based on Personal Computers (PCs). For each PC a vocal interface has been implemented. It allows a disabled user to perform vocal commands through a phonetic terminal linked to it. In addition, it also allows a vocal browsing and querying by a module interfaced to a computer and, by a modem, to a phone socket.

From the software point of view, the system consists of a tool able to understand a pre-defined language, which is a customizable set of commands, somewhat like a template, and which can be issued through the voice in different applications. The system contains some "voice files", which contain the recordings of a person's voice saying specific commands (such as "new", " page previous", etc.) or keywords (such as "phone agenda", or "one", etc.). These files or commands are used by the tool to execute the associated commands in a corresponding internal language.

By this vocal interface it is possible to implement and manage:

1. access to the communications network (e. g.; Internet);
2. local activities (e. g.; word or excel);
3. activity of Teleworking and Teleteaching.

They form the principal functionalities of the implemented system which is described in the following section.

3 The Project

The temporal phases of the project are the following:

1) definition of the qualifications necessary to the user;
2) definition and project of the vocal recognition system;
3) implementation of the above-mentioned system and of the relative man-machine interface;

4) use of the system by selected disabled people and debugging of the system.

In the following, we briefly discuss on the first three points.

3.1. Definition of the necessary qualifications for the user

The first step of the project consists in the definition of the necessary qualifications for the user. This fact is linked to the type of use of the system which the user wishes.

We have defined a subset of the user necessity which regard to simple operations (but for which it is necessary the use of hands), or tele-access to software programs (such as, the tele-access to a database or the use of commercial software), or the telework (this last one linked to both the previous points). We also foresee its use for developing a thesis in a university environment.

3.2. Definition and design of the vocal recognition system

The second step of the project consists in the definition and design of the vocal recognition system. The project, known the goal, consists of the following modules:

- a tool for transforming the voice in commands for the computer;
- a language for the application (that is, a list of structured vocal commands);
- a database in which the user can browse in order to choose the function(s) which he/she wishes to activate;
- a user friendly vocal interface for the man machine communication.

3.3. Implementation of the free hand system

The third step of the project consists in implementing the free hand system described as the functionality of the system.

4. The System Functionality

The functionality of the implemented system consists of three different activities. Each type of these activities presents the potentiality of the above mentioned system as a support to disabled users for performing the desired action.

4.1. Management of access modality to the communication networks

The functions which can be activated by this type of management are the following: a) activation of communication; b) execution of communication; c) disactivation of communication.

In this phase, particular attention has been paid to "Internet". Network navigation through common browsers (i.e., Netsacpe or Explorer) is charaterized by small movements (i.e., placing the mouse on the chosen voice) which in some cases are very difficult or almost impossible, as it is in the case where users have limited motory abilities or have an imprecise functioning mouse. In the case in which the mouse is guided by a vocal identifier (for those unable to use their hands) things are even more complicated.

To solve this problem, the idea is to have a screening of the document on the network before it is visualized by the browser, select all links present (those that allows us to address other documents) and visualize them in a specific list so that the user can easily utilize them.

Therefore, a specific provider called *Winsock2 Hook provider* for monitoring and filtering the network traffic has been designed and implemented. Through the use of *Winsock2 Hook provider* a new application called *Link Extractor* has been created. It allows the motory disabled to navigate in the World Wide Web (WWW) in an easy manner through the creation of a link list easily available by voice through a vocal interface (see Figure 3). To each item of this menu a number or a letter is associated, since they can be easily pronounced by a disabled user . In this way, a disabled user can make a "vocal" web surfing.

The *Link Extractor* is characterized by three modules having three activities:

- intercept network traffic by means of hook functions retrieved by the *Winsock2 Hook provider*;
- provide a cache mechanism alternative to the browser, since it is necesssary to discativate the internal cache in order to intercept all request coming from the browser;
- visualize the list of links selected from the document with a visual object integrated with the browser itself.

To begin with the *Link Extractor* can be seen as an extension of the *Winsock 2 Hook Provider*, from which it is charged as *DLL* (Dynamic Link Library) *WinSock2*. The architecture of *Winsock2 Hook Provider* is shown in Figure 2.

The API (Application Provider Interfaces) of *Winsock2 Hook provider* provides a simple interface for configuration and for monitoring the network activities. The monitoring of network has been managed by pre-stabilised hook functions which should be exported by each monitor, implemented as DLL. These functions will be retrieved by *Winsock2 Hook provider*. Instead, the configuration has been managed by some functions provided in the above mentioned provider which allows to register DLL monitors. Thanks to this potentially, *Winsock2 Hook provider* represents an optimal instrument in all of these cases which are necessary to extend the functionality of a provider.

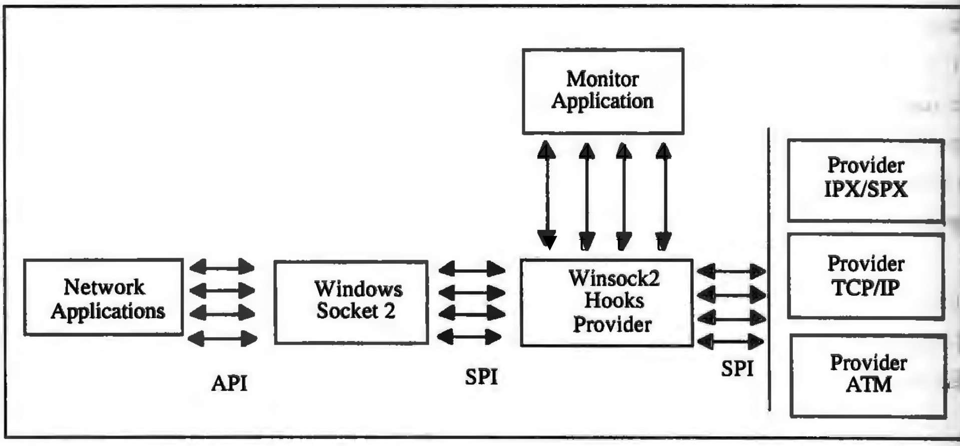


Figure 2. Architecture of Winsock2 Hook Provider

As mentioned above, it allows to manage different network protocols by means of a unique application interface for the configuration and management of monitoring requests. Through this interface every application can indicate a series of hook functions which *Winsock2 Hook provider* can perform to deviate the network traffic on the same applications.

Each application which desires to display the flow of data on the network must have the following hook functions available for the *Winsock2 Hook provider*.

a- *Hook before reception.* It will be invoked every time the network is about to reach the waiting state for data. The monitoring application prevents this operation from being carried out or simulates its completion.

b- *Hook after reception.* It will be invoked every time the network has completed an reception operation. The received data will be available for hook before being available to the network that manage this operation.

c- *Hook before transmission.* It will be invoked each time the network approaches a transmission of data. The monitoring application can modify the contents of the transmitted data or prevent totally the trensmission.

d- *Notification of interrupted connection.* This notification makes the monitoring systems independent from the type of transmission.

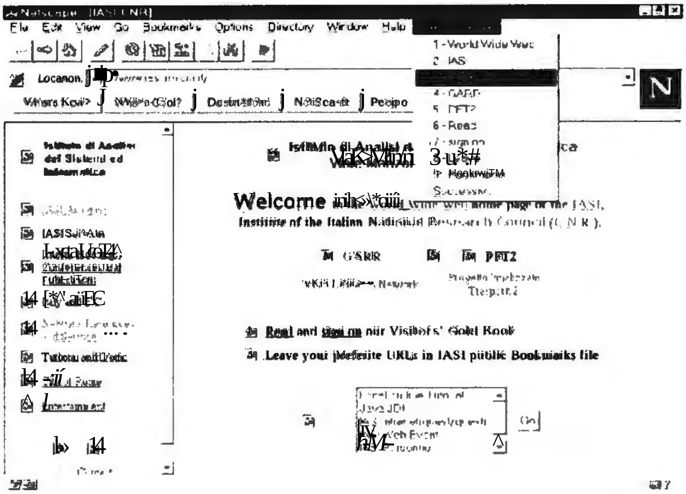


Figure 3. The menu of links related to a web page

The principal functionality of *Winsock2 Hook provider* consists of providing a versatile support independent from the type of protocol used to extend the functionality of the network. Table 1 summarises all actions that are possible to undertake using *Hook provider* and those that require a direct interaction with the network.

Allowed actions	Non-allowed actions
• Intercept and modify transiting data on network • Determine the destination of transmitted data • Determine the source of received data • Prevent a transmission from being finished • Prevent a local application from receiving data • Simulate the errors of network • Deviate the transmissions on different addresses • Simulate particular conditions of network (events) such as reception of data, etc.	• Throw down all stabilised connections from final applications • Carry out network interaction different from those stabilised from final applications • Carry out any network activity which is not a simple modification of transmitted and received flow data from final application • Monitoring the Quality of Service

Table 1. Allowed and non-allowed actions

This provider is not an implementation of a particular protocol and it is not a support to develop the network applications. Its unique task is to easily provide data transiting on network in output. In effect, it origins from the extension of HTML browser normal functionality to be adequate for the disabled user, in particular in the case of motory disabled.

4.2. Management of local activity

The local activity management allows a disabled user to have, by a set of implemented macros, a normal use of the computer (for example, writing and printing a file text, etc.) through vocal commands.

In this way, every time the user has to select an action and to carry out a given action (for example, writing an equation), it is sufficient that the user says the name of a given symbol to which the implemented macro is associated. In Figure 2, an example is shown: the user pronounces "Integral" for visualizing the relative symbol in the Equation Editor's window (see Figure 4).

It is important to note that the system answers to *sounds*, not to *meanings*.
If the user says the number corresponding to a given icon (and, then, to a given function) which he/she wishes to activate, this action selects the associated button and, through the relative key-word, the system emulates the click on the icon, so that a new configuration appears on the screen.

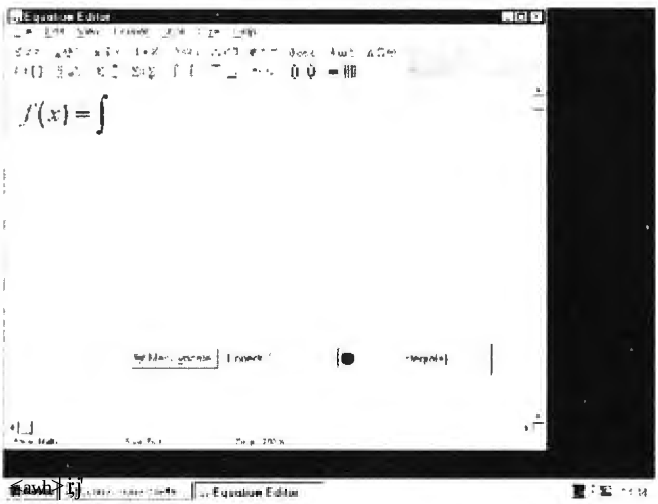


Figure 4. Example of a local activity

A simple vocal language is created in order to browse among the different levels and along the same level in which a given tool is generally divided. For example, to change level the user can follow the previous procedure (saying the number near the icon) or, to go back (saying the key-word "close"). Instead, to browse inside the same level, the user uses the key-words "go", "back", "next" and "previous".

This kind of facility is provided for each type of application. In this way, the disabled user can carry out its own activity, through vocal training of the implemented macros.

4.3 Management of Teleworking and Teleteaching activity

In the context of the proposed system, two important types of activity by means of Telematic services for disabled users have been identified. They are teleteaching and teleworking, still in progress. As for teleteaching, it will be possible for a disabled user to put into action either an academic activity (for example; preparing a theses at a distance) or a professional formation activity (for example; teaching/learning a programming language). While for teleworking, we have considered two different scenarios. The first one is relative to a typical office work, the second refers to learning those modalities about the navigation on a Telematic network and accesses to multimedia databases.

5 Conclusions

The project" Help communication systems for the disabled " and its sub theme "Free hand access system to the telecommunication and the telematic services by vocal commands" has, as a goal, the implementation of a complex system, able to help this type of disabled people in different activities through the use of a computer. In such a way, a disabled user carry a normal activity (with vocal commands associated to apposite macros) using all the applications installed in a computer. The user can also have an HTTP access to a WWW server through HTML pages in "Internet" and navigate in it by saying a simple key-word (a number or a letter) which corresponds to a command (or an icon) organised by the system in an additional menu in the web page. Finally, the system allows the user to carry out a teleworking and teleteaching activity in order to release them from a fixed location when he/she wishes to use computers.

6 References

[1] RUTTER, D.R., Stephenson, G.M., Dewey, M. E., Visual communication and the content and style of conversation, British Journal of Social Psychology, 20(1981).

[2] World Health Organization, International classification of impairments, disabilities, and handicaps. WHO, Geneva 1980.

[3] FREDERIKSEN, J., MARTIN, M., PUIG DE LA BELLACASA, R., VON TETZCHNER S., Use of telecommunication: The needs of people with disabilities. Madrid: Fundesco, 1989.

[4] BOWE, F.G., Disabled and elderly people in the first, second and third world, in: International Journal of Rehabilitational Research, 13 (1990).

[5] 4th International Conference on Computers for Handicapped Persons, in: Lecture Notes in Computer Science, Springer-Verlag Publication., 860, Vienna 1994.

[6] Telematics for Health and Disabled & Elderly People, in; Joint Conference and Exhibition of the European Union and South Africa, Gauteng, South Africa 1996.

[7] Interdisciplinary Aspects on Computers Helping People with Special Needs, in: 5th International Conference, ICCHP '96, Linz, Austria 1996. Interdisciplinary Aspects on Computers Helping People with Special Needs, in: 5th International Conference, ICCHP '96, Linz, Austria 1996.

[8] RAFANELLI, M., Proposal of a system for the free hand access to telematic services by vocal commands, in: International Conference on Health Telematics, Ischia, Italy 1995.

[9] Telemedicina-Tema 8: Sistemi di ausilio per la comunicazione dei disabili. Ministero dell'Universita e della Ricerca Scientifica e Tecnologica, 1994.

[10] Telemedicina - Sottotema 4 (tema 8): Sistema per l'accesso a mani libere ai servizi di telecomunicazioni e telematici tramite comandi vocali, 1994.

A Chinese Text-to-Speech System for People with Disabilities [1]

Li Jianmin, Chen Weijun, Yang Jinfang, Lin Fuzong, Zhang Bo

State Key Laboratory of Intelligent Technology and Systems

Department of Computer Science, Tsinghua University

Beijing(100084), China

Abstract

This paper presents a computer system designed for disabled people to help them communicate with others like normal people. One of its main subsystems, a Chinese text-to-speech (TTS) system, is then introduced in detail. Some problems in TTS system are also discussed.

1 Introduction

Every day, many people unfortunately become disabled because of all kinds of inborn or postnatal reasons. Disability can be divided into three types: activity disability, such as lameness; communication disability with environment, such as blindness, deafness and dumbness; cogitate disability. Any disability will bring them great inconvenience in daily life and numinous pain. Today, it is especially difficult for those who lost ability of communication partly or fully to live in such an information society.

Generally, we get information of external world through different media, such as graphics, images, and sounds, by sense of visual, hearing and touch. We also deliver information to outside through

[1] supported by National Nature Science Foundation of China

articulatory organ. So barrier of vision and hearing will break off the way of getting information dysphonia will block the output of the information. Scientists have tried to find other alternative for these people to communicate with environment for a long time. In fact, some methods have been found, such as Braille for the blind, sign language for the deaf and the dumb. But these methods are quite complex and not convenient to be mastered and used because they are not the nature way for people to communicate. Moreover, there are still some open problems, such as communication between these people with disabilities and those normal people, communication between these people. It is impossible for them to communicate with each other without persons who master several methods. And the people with disabilities can not make use of things of normal people, such as newspaper, book, and broadcast. This makes these people disadvantageous position in the information society.

Nowadays, with the development of science and technology and popularization of computer, some more convenient and natural methods are provided for these people. A lot of products with speaking prompt for the blind have been put into the market, including watches, reading machines, etc. Furthermore, with the help of computer, communication between the normal people and the people with disabilities and communication between the people with disabilities becomes easier. The people with disabilities can live a normal life to a large extent.

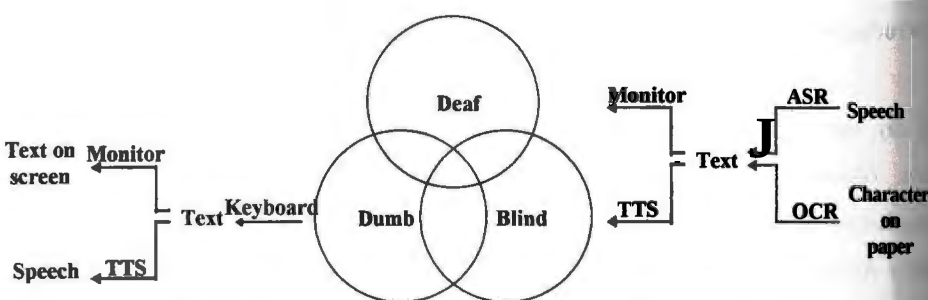


Figure 1. How the computer system help the people with disabilities in communication

As shown in *Figure 1.*, the computer system can transform speech of outside to text by automatic speech recognition (ASR) for people who are only deaf to read on the screen of the computer. The system can also convert the characters on the paper to text by optical character recognition (OCR), then convert text to speech through text-to-speech (TTS), for those who are only blind. Even for those who are both deaf and blind, it is possible to convert the text to Braille. The dumb can "talk" with others through keyboard, and the text they input can also be converted to speech by TTS, so they can even talk with others like normal people with the help of the system.

Such system is no longer just an imagination. Scientists have made such great progress in the fields of ASR, OCR and TTS that systems are more and more practicable. Research on Chinese transformation system is undergoing in our lab. The system architecture is as follows:

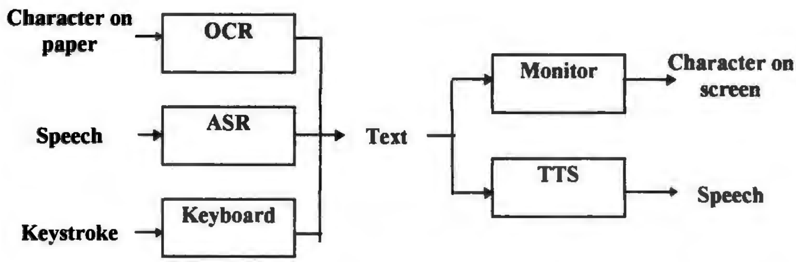


Figure 2. System architecture

Recently, we have constructed the whole system and are going to improve its performance. The task of our group is to construct a text-to-speech subsystem. In the following, the architecture and some technical problems of the TTS are discussed in detail.

1.2 Key problems in TTS system

A TTS system must solve two main problems. One is accuracy of the speech, the other is naturalness of the speech. Although TTS systems today have achieved a high level of accuracy and intelligibility, their voice quality is still quite poor. There are mainly two reasons:

Firstly, the system can not get enough prosodic information during the stage of text analysis. The written form of a sentence only specifies the words to be read. The pronunciation of each word, the information of prosody, including which word should be accent, the duration of each word, and intonation of the sentence, must be determined by the reader (human or machine). In Chinese, there are some special problems. For example, since a Chinese word consists of different numbers of Chinese characters, and its boundary in a sentence is not delimited by blanks, the boundary should be decided by the reader at first. This procedure is called word segmentation. In ancient Chinese, there are even no punctuations. Most people can deal with these tasks easily and perfectly with the help of their understanding to the sentence and their experience and habit, besides the lexical and grammar knowledge. But TTS system does not. It is not strange that the synthesis voice quality is not good because the system is short of essential information. In fact, a computer does not understand what it is reading at all.

Secondly, the synthesizer is not perfect. Currently, the methods of speech generation can be classified into two categories according to different representation of speech signal. One tries to explain the generation of speech signal and simulates this procedure, while the other tries to find out efficient encoding schemes for speech signal which are also appropriate for prosody control. There are two methods of speech modeling, articulatory synthesizer and formant synthesizer. At present, articulatory synthesizer is rather far from applications and marketing because of its computation cost and the unsolved underlying theoretical and practical problems. Formant synthesizer can produce good quality speech, if the appropriate parameters are obtained. It needs lower storage than encoding methods and less computation than articulatory synthesizer, but the appropriate parameters are not easy to generate. There are a lot of encoding methods for speech signal, such as pulse code modulation (PCM), linear predictive coding (LPC), code excited linear predictive coding (CELP), etc. With the help of pitch-synchronous overlap-add (PSOLA) method, the encoding methods are more and more popular now. The synthesis voice is very satisfying while the storage required is quite large compared with other methods.

3. Our TTS system

Our TTS system can be divided into seven modules, as shown in *Figure 3*.

3.1. Text preprocessing

Non-Chinese characters in the original sentence such as number, symbol, foreign language word are first converted to the corresponding Chinese characters.

3.2. Word Segmentation

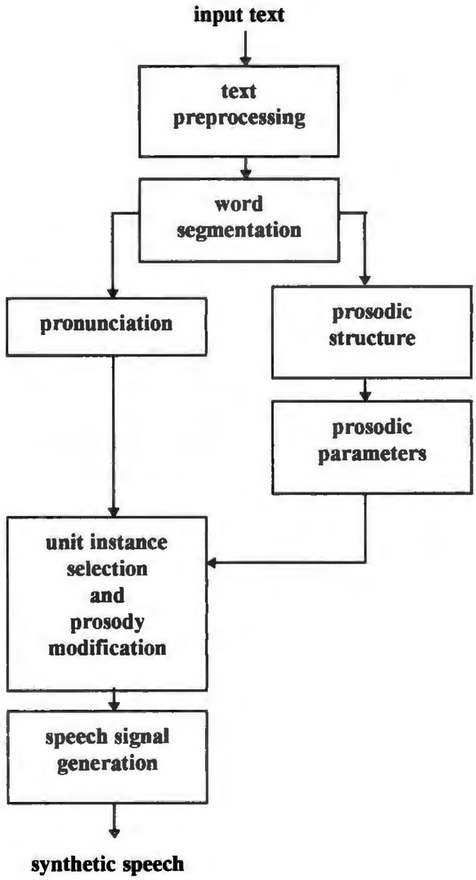


Figure 3. The block diagram of our TTS system

In Chinese, there are about 10000 characters and more than 50000 everyday words which consist of several characters. Although most characters have their own pronunciations and meanings, the basic unit for Chinese linguistic processing is the word. For example, about 1/8 of the Chinese characters have more than one pronunciation, so the appropriate pronunciation must be chosen according to the particular word that the character belongs to. As mentioned earlier, however, the words' boundaries in a sentence are not delimited by blanks. So it is necessary to separate the text into words first after preprocessing.

Word segmentation is always a hard problem in Chinese language processing. The simplest way is to match the word in a very large dictionary that enumerates the whole Chinese words. But sometimes it does not work because the input text has more than one possible segmentation. Both heuristic and statistic methods can be used to choose the best one. The most common heuristic method is maximum matching (MM) that finds the word with the most characters at any point from beginning to end of the sentence (forward MM) or from end to beginning of sentence (backward MM). This method can not solve the problem successfully because there are a lot of words that do not exist in the dictionary, such as person name, toponym, derivative. Thus nowadays, the statistic methods which calculate the most probable sequence of words according to some models become more and more popular.

However, it is found that reading the words with longer than three characters one usually divides them into several subwords that consist of one or two characters rhythmically. This phenomenon indicates that the phraseological words and the rhythmical words perhaps are not the same and that word segmentation in a TTS system are not as strict as in natural language processing (NLP). It is observed that very few wrong pronunciations resulting from wrong word segmentation do not effect the naturalness of the whole sentence too much. Thus, in our TTS system, the simple MM method is employed. Since most of Chinese words are composed of two to four characters, the dictionary in our system only consists of these words. In order to improve the rhythm, we also combine two or three isolated characters into one word. In fact, the result is quite satisfying.

3.3. Pronunciation

Once the segmentation is completed, pronunciation of each character can be decided through a dictionary according to the word it belongs to. In fact, in our system, the dictionary is divided into two parts in order to save the storage and improve the efficiency. Only the words that contain at least one character having more than one pronunciation are stored in the first part. The second part consists of the whole Chinese characters and their pronunciation. Different pronunciations of the same character are ordered according to the frequency of the pronunciation. The pronunciation of each character in the whole word is assigned to the first pronunciation of this character in the second part of the dictionary. If there is a character having more than one pronunciation in a word, and if this word can be found in the first part of the dictionary, the whole word's pronunciation is changed according to the item in that part.

3.4. Prosodic structure

At present, an important and difficult problem in TTS system is to generate natural prosody from the plain text. Prosodic structure must be obtained from the text at first in order to break the sentence into several intonational phrases. It is different from grammar structure. It makes the text easily to be read and more natural. A method introduced by AT&T, in which a decision tree model that is trained on a corpus of text annotated with prosodic phrase-boundary information is used to predict the prosodic structure [5].

1.3.5 Prosodic parameters

Prosodic parameters including pitch, duration and energy are then calculated on different levels of the prosodic structure. It is pointed that stress is an important parameter which connects pitch, duration and energy in Mandarin Chinese [2]. So, stress degree of each syllable is first determined in the whole sentence according to some rules. And then, the fundamental frequency (FO), duration and energy of each syllable are calculated from their stress degree.

1.3.6 Unit instance selection and prosody modification

In our system, concatenative waveform synthesis, which is a kind of encoding methods, is employed. Compared with other generating methods, it can generate more natural speech unit since the explicit speech model is avoided and all the details of the unit are reserved. But even though the synthesis units in the speech database are selected carefully, the performance is not good, if we just concatenate several units together. The reason is that the context of the unit in the stored speech unit database, including phonetic and prosody, is different from the context in the synthesized sentence.

Unit selection method is used to improve the naturalness of the concatenative speech synthesis [3]. Many instances of the same synthesis unit are stored in the speech database. The most appropriate instance is selected from the database according to the context. Although the best instance sequence of synthesis unit can be found, the context of each instance may be not the same as what we want, due to the limitation of the database.

Another improvement to the concatenative speech synthesis is to modify the speech unit according to prosody that we need. The most popular method is PSOLA, including time domain PSOLA (TD-PSOLA), LPC-PSOLA, etc. [4] Although they can adjust the supra-segment feature of the unit in time domain, such as pitch, duration, and energy, there are some disadvantages with such an approach. Firstly, the segment feature can not be modified. Secondly, the extent of prosody modification can not be too large, otherwise there will be great distort.

In our TTS system, two approaches are combined in three ways. The first one, and also the simplest one, is that after the best instance sequence is selected, prosody modification is applied to each instance according to distance between the target context and the original context. The second one is that after the best instance sequence is selected, the best extent of prosody modification is also calculated. The third one is that the best instance sequence and the best extent of prosody modification are computed in one step by calculating an improved cost function. It is clear that if each unit only has one instance, this method is degenerated to the second approach and that if no prosody modification is done, this method is again degenerated to the ordinary unit selection approach. Among these three methods, the first one is the simplest one, but the naturalness is inferior to others. And the third one is the most complex one, while the naturalness is the best.

3.7. Speech signal generation

This module generates the speech signal from the input parameters. It can be implemented by hardware, or simulated by software. As mentioned earlier, it can be an articulatory synthesizer, a formant synthesizer or a coder synthesizer. In our system, we use a simple PCM coder synthesizer which just transfers the input data to the speaker.

4. Conclusions

In this paper, we have outlined a computer system for people with disabilities to communicate with others. A Chinese TTS system which can convert Chinese text into speech is then introduced in detail. We will continue to improve the naturalness of the system. With the improvement of the system's performance, the TTS system can bring more convenience to the people with disabilities.

5 References

[1] Richard W. Sproat, Joseph P. Olive, "Text-to-Speech Synthesis", AT&T TECHNICAL JOURNAL, MARCH/APRIL 1995, pp. 38

[2] Xingjun Yang, Huisheng Chi, etc., "Speech signal digital processing", pp. 319

[3] Andrew J. Hunt and Alan W. Black, UNIT SELECTION IN A CONCETENATIVE SPEECH SYNTHESIS SYSTEM USING A LARGE SPEECH DATABASE, ICASSP 1996

[4] Moulines E and Charpentier F, "Pitch synchronous waveform processing techniques for text-to-speech synthesis using diphones", Speech Communications, 9 (1990), pp. 453-467

[5] M. Wang and J. Hirschberg, "Automatic classification of intonational phrase boundaries", Computer Speech and Language, Vol. 6, No. 2, 1992, pp. 175-196

Current Developments in Low-Budget Speech Recognition Systems

Jill Hewitt

Abstract

This paper explores the requirements for a general purpose speech recognition system for use at the university by people who have difficulty using a keyboard or prefer not to use one. It investigates the "out of the box" claims of three low cost recognition systems, reflects on the improvements in speech recognition over the past few years and makes recommendations for their use.

1 Introduction

At last there are a range of affordable speech recognition products on the market which will run on "normal" PCs – at least if not on the very basic machines, they will run, without the need for any specialist hardware, on the type of multimedia PC that many people are purchasing today. What is more, the current products are cheap – all of the products mentioned in this paper are under £200, two were under £50.

In this paper we identify a set of requirements for a speech recognition system to be used at the university by people who have difficulty using a keyboard, or who choose not to use one. We consider five of the most common speech recognition products and evaluate the "out-of-the-box" performance of three of them in terms of recognition accuracy, speed of text production and ease of use. We identify their good and bad points and assess the extent to which these products might meet our requirements. Finally we reflect on the advances made in speech recognition technology over the past few years, identify lessons learned from the product evaluations and make recommendations for further improvements in the products.

2 Requirements for a Speech Recognition System

The purchase of off-the-shelf products is often a hazardous affair; we have to rely on information from company advertising and press releases, and often on what is written on the box. It is all too easy to be dazzled by the marketing claims or to overlook some vital information which shows that the product will not be suitable for our needs. The problem of acquiring requirements for off-the-shelf products has recently been addressed by Maiden and Ncube [3] who propose a new method, PORE (Procurement-Oriented Requirements Engineering) which at the time of writing is still under development. This uses a set of templates to guide the requirements engineer through the process of obtaining customer requirements. The first template is based on using supplier data, the second on a product demonstration and the third on a hands-on product evaluation; at each stage, the choice of products can be restricted to those which best match the customer requirements and the customer requirements can be refined in the light of the customer's increasing knowledge of the products. The method is perhaps targeted at more complex scenarios than the one in hand, but there are nevertheless some useful lessons to be learned. We began by considering our requirements in the light of information provided by suppliers and the manufacturers' web sites and we identified both essential and preferred requirements. We distinguish between requirements for keyboard and non-keyboard users since it is generally accepted that it is more efficient to use multi-modal input, for example for editing and moving around a document, *if* the user is capable of doing so (see for example [1] and [2]). We have been able to amalgamate the process for templates 2 and 3 in the case of the very cheap products, since we have dispensed with the product demonstration in favour of buying the product for a hands-on evaluation – not a course we could have followed a few years ago when products like these cost in excess of £1,000!

2.1 Essential Requirements

We are looking for a speech recognition system which can be used by students and/ or staff at the University to carry out tasks on a PC. The author is often asked to recommend products for disabled students who have limited or no use of the keyboard, and, increasingly, there is concern amongst staff who use a keyboard continuously that they may suffer from RSI. Some people, may simply prefer to use speech rather than a keyboard. There is therefore an eagerness amongst people to try speech input as a viable alternative to the keyboard for at least part of the time. The number of tasks people may want to carry out is varied, but all users need to be able to produce documents in Microsoft Word which is in common use at the university. Our essential requirements for a speech recognition system are as follows:

Reqt. No.	Requirement	kbd user	non-kbd user
R1	It must work with our 'standard' multimedia PCs and Windows 95 i.e. 16MB RAM	✓	✓
R2	It must be easy to learn and use 'out of the box' without the need for specialist training	✓	✓
R3	It must be capable of producing documents in Microsoft Word or documents that can be translated into Microsoft Word.	✓	✓
R4	It must produce text at least as fast as other means at the user's disposal (e.g. switch input, special entry pads)		✓
R5	It must be capable of producing text as least as fast as an unskilled typist.	✓	
R6	Once set up, it must be capable of being used without recourse to the keyboard for error correction.		✓

2.2 Preferred Requirements

In addition to the essential requirements, a number of other requirements were deemed desirable:

Reqt. No.	Requirement	kbd user	non-kbd user
R7	The entry-level system must be low-cost – preferably under £100.	✓	✓
R8	It should be compatible with other Microsoft products e.g. MSAccess and Excel	✓	✓
R9	It should be capable of controlling the functionality of Windows 95		✓
R10	It should be able to support more than one user	✓	✓

3 The Products

Five products were chosen as a representative sample of affordable speech recognition systems which are currently available 'off-the-shelf' in the UK. These are all marketed as 'entry level' systems with further features being provided in add-ins or in more sophisticated versions of the product. Three of the products provide 'discrete word' recognition – that is, the user has to leave a small gap between words when dictating; the other two provide 'continuous speech' recognition. All of the products are 'speaker-dependent', that is they adapt to the user's voice with the result that recognition should improve over time. Each product is supplied with its own microphone. A brief specification of each product is given below, details are taken from the product box or from the suppliers web site. The price given for the first four is the actual purchase price we paid, we were unable to purchase the fifth product in time for this evaluation:

Name: DragonDictate® Solo **Price:** £44
Supplier: Endeavour Technologies [4] **Manufacturer:** Dragon Systems UK Ltd [5]
Purpose: "Dictate to your PC and control your Windows applications by voice alone" .... "no typing!"
Platform: Windows 95 and Windows 3.1
Applications: Microsoft Word™, Excel™, Powerpoint™, Access®, Internet Explorer
Type of recognition: not explicitly stated on the box but discrete word utterances is implied: "Dragon's QuickTalk™ technology makes dictation faster than ever before by shortening the pause required between words"
Vocabulary: UK English dictionary and language model, 120,000 word vocabulary with 10,000 word active customisable dictionary. (also available for American English, French, German, Italian, Spanish and Swedish.
System Requirements: 486/66 MHz IBM-compatible PC, Pentium processor recommended, Industry standard 16-bit sound card or built-in audio systems, including the SoundBlaster 16®, 37MB hard disk space, CD-ROM drive, Mouse and keyboard recommended, Minimum 16MB memory, recommended 20MB.

Name: VoiceType® Simply Speaking **Price:** £49
Supplier/Manufacturer: IBM [6]
Purpose: "Turn your spoken words into... notes, papers, and e-mail by simply speaking to your computer"
Platform: Windows 95
Applications: uses own 'VoicePad' for dictation
Type of recognition: discrete word – "Simply Speaking recognises "discrete" speech patterns – that is, speaking with a very slight pause between words"
Vocabulary: UK English dictionary and language model, 64,000 word vocabulary (also available for American English, French, German, Italian, Japanese and Spanish)
System Requirements: processor performance equivalent to Intel Pentium 100MHz or faster IBM-compatible PC, Industry standard 16-bit sound card - Creative Labs SoundBlaster 16® or 100% compatible or IBM Mwave sound card, 38MB hard disk space, CD-ROM drive, Minimum 16MB memory, recommended additional RAM when using with word processors such as MSWord, Lotus, WordPro and Corel WordPerfect.

Name: Kurzweil VoicePro **Price:** £149
Supplier: L&H Direct [7] **Manufacturer:** AlphaSoftware [8], Kurzweil AI[9]
Purpose: "Create documents, spreadsheets, e-mails and more simply by speaking!". ".... works with virtually all of your Windows applications"
Platform: Windows 95, Windows 3.1
Applications: Microsoft Office 97 and Lotus SmartSuite 97
Type of recognition: discrete word – "Kurzweil VoicePro™ allows you to control your Windows applications; creating memos, e-mails, reports, and presentations through discrete speech" plus a "Continuous Digit Recognizer"
Vocabulary: 60,000 word active vocabulary, 200,000 word dictionary. American English only.
System Requirements: Pentium processor for Windows 95, 486/DX4 75 for Windows 3.1x, Soundblaster 16-compatible sound board, or other Windows-compatible 16-bit sound board, 35MB hard disk space, 24MB RAM

Name: ViaVoice™ **Price:** £99
Supplier/Manufacturer: IBM [6]
Purpose: "Talk to your computer. Easy to use, speak normally. Dictate directly into Microsoft® Word"
Platform: Windows® 95 , Windows NT® 4.0
Applications: uses own 'VoicePad' for dictation or Microsoft® Word
Type of recognition: continuous speech – "ViaVoice uses continuous speech technology so you can speak without pausing between words"
Vocabulary: UK English dictionary and language model, 30,000 word base vocabulary , expandable to 64,000 words (also available for American English, French, German, Italian, Japanese and Spanish.
System Requirements: 166MHz Pentium® processor or faster, Industry standard 16-bit sound card - Creative Labs SoundBlaster 16® or 100% compatible or IBM Mwave sound card, 120MB hard disk space, CD-ROM drive, 32MB memory for Windows 95, 48MB for Windows NT 4.0.

Name: SoloNaturallySpeaking Price: £99
Supplier: Endeavour Technologies [4] Manufacturer: Dragon Systems UK Ltd [5]
Purpose: "Produce reports, proposals, letters, e-mail and more in one half the time it takes to type"
Platform: Windows® 95, Windows NT 4.0
Applications: Uses own window for dictation, compatible with 'common everyday applications'
Type of recognition: continuous speech "Speak to your computer naturally and at a normal pace – up to 160 words per minute and more – without pausing between words"
Vocabulary: UK English dictionary and language model, 230,000+ word vocabulary with 30,000 word active vocabulary, (also available for American English)
System Requirements: Minimum 133MHz Pentium Processor IBM-Compatible PC (faster with MMX), Industry standard 16-bit sound card or built-in audio systems, including the SoundBlaster 16®, 60MB hard disk, CD-ROM drive, 32MB memory for Windows 95, 48MB for Windows NT.

4 Compatibility with our Requirements

Having gained the technical data for each product we were in a position to compare this with our initial set of requirements, following the recommendations given in [3]. The table below sets out our findings:

Reqt	DragonDictate® Solo	VoiceType® Simply Speaking	Kurzweil VoicePro	ViaVoice™	SoloNaturally Speaking
R1	Yes	Yes	not all	not all	not all
R2	Apparently	Apparently	Apparently	Apparently	Apparently
R3	Yes	Yes	Yes	Yes	Yes
R4	Apparently	Apparently	Apparently	Apparently	Apparently
R5	Apparently	Apparently	Apparently	Apparently	Apparently
R6	Yes	No	Yes	No	No
R7	Yes	Yes	No	Yes	Yes
R8	Yes	No	Yes	No	No
R9	Yes	No	Yes	No	No
R10	Yes	Yes	Yes	Yes	No

At this point in our investigations it appeared that DragonDictate®Solo most closely met our essential requirements (R1 to R6) as well as our preferred requirements. It was apparent that only two of our products met R6 and would be suitable for non-keyboard users and only two of them would work on our lowest specification computers. We were still unsure about the ease of use of the products (R2) and their speed (R4 and R5), and we could have sought out product demonstrations at this point. However, as we were able to purchase four of the products we were able to proceed directly to hands-on evaluations. Additional information was obtained from [10] which compares ViaVoice and Solo NaturallySpeaking, praising both products and coming down slightly in favour of NaturallySpeaking. We concluded from this article that NaturallySpeaking would definitely be worth evaluating and propose to include it in our longer term trials. The choice between discrete word or continuous speech recognition had not been raised in our requirement*] mainly because we had no prior experience of using a continuous speech system; the hands-on evaluation would give us a better idea of their potential.

5 Product Evaluations

The product evaluations were planned in two stages. The first stage, which is reported here, was to test the 'out-of-the-box' performance of the products with a number of potential users. The second stage is a longer term trial over two months where each system will be used by a volunteer to carry out part of their daily work.

5.1 First Stage Evaluations

Each of the four systems to be tested provided instructions for getting started which would enable the user to set up the system and undertake an initial dictation session without the need for complicated training sessions (although, to be fair, the manufacturers do point out that recognition is likely to improve if the user invests time in 'enrolling' their voice by speaking a number of pre-defined words or phrases which the system can use to learn the user's voice patterns).

In a previous evaluation [11] carried out in 1994, we attempted a similar exercise with very disappointing results and we were interested to see how much speech recognition had improved in four years. In the 1994 trial we tested an expert, a novice and a new user on hands-free use of the DragonDictate system, getting them to input the same text and counting the error rates and speed of input The results were as follows:

Expert: 38 words per minute, a recognition rate of 91%
Novice: 8.5 words per minute, a recognition rate of 67%
New User: 3 words per minute, a recognition rate of 34%

We concluded from this that the initial poor performance of the system might be enough to dissuade users from persevering with its use unless they had no alternative but to use it. In a questionnaire received from 20 (then) recent purchasers of the DragonDictate and IBM VoiceType systems we found some degree of dissatisfaction amongst users who had needed an average of one month to get used to their system (and this after receiving typically half to one day's training from a consultant).

5.1.1 The Evaluation

This involved seven users, all native speakers of UK English, one of whom was the new user in the previous study and three of whom (including the 'new user') had some previous experience of speech recognition systems. All of the users were experienced keyboard users, but three had experienced RSI-type problems and were keen to investigate alternatives to keyboard input. Since not all of the systems were capable of being used in totally 'hands-free' mode we decided to test them all in the environment where a keyboard could be used to input corrections.

Our original intention was to test all of the users with all four systems, but two problems made this impractical. The first was that we quickly discovered that the out-of-the-box performance of the Kurzweil VoicePro was unacceptably low – we were getting a recognition rate of only 50%! After consultation with the supplier, we confirmed that this is because the VoicePro, unlike the other products, was not supplied with a UK English language model, in order to get acceptable recognition rates we would have to undertake the lengthy enrolment process. We decided therefore to leave VoicePro out of our stage one evaluation though we will include it in our longer term trials as it has received favourable press reports. The second problem was one of time and commitment - users could not evaluate all three products in the allotted 90 minute sessions. We settled on a compromise – four users evaluated the two discrete word systems DragonDictate Solo and Simply Speaking and three users evaluated a discrete word system, DragonDictate Solo and the continuous speech system, ViaVoice.

In each case, the system was pre-installed for the user but s/he was asked to follow the manufacturer's instructions for setting up the microphone and carrying out all the necessary steps for making a first attempt at dictation. The users were then asked to input a standard text which comprised 100 words (including punctuation) and follow the instructions for correcting any errors. They repeated the text two more times, each time correcting the recognition errors, prompts were given by the evaluator if they got really lost and notes were taken during and after each session. The tasks involved for each of the systems prior to the dictation test were as follows:

DragonDictateSolo®

- Create a new user
- Set up and adjust microphone
- Run Quick Enrolment (less than 20 words)
- Run the first four lessons of the on-line tutorial
- Bring up Word Pad by voice

Simply Speaking

- (no need to create a new user as this is done when the system is installed – for the purpose of the evaluation each new user was created prior to the test)
- Set up and adjust microphone
- Follow the recommendations from the manual on 'Learning How to Dictate to IBM Voice Pad' – this involves reading selected parts of the on-line Help system and dictating a sample sentence

ViaVoice

As this is an upgrade to Simply Speaking the instructions for getting started are almost exactly the same, with the additional requirement that the user must dictate 14 sentences into the Dictation Trainer before starting. Dictation can be made either into the SpeakPad or directly into Microsoft Word. For this exercise we used SpeakPad.

DragonDictate Solo works in a different way from the two IBM systems. The user corrects errors as s/he goes along. A list of possible words is presented to the user. If the word they said is at the top of the list then they need do nothing, if it is in the list they can say "Choose n" to select the nth word in the list, if it is not in the list they can start typing (or spelling if they are in hands-free mode) until it appears in the list and then say "Choose n". The command "Scratch That" is used to delete the previous word and the command "Oops" brings up the word history so that words up to 12 words back from the current word can be corrected. The details of error correction are described in the tutorial and the user is given the chance to practice. The IBM systems on the other hand assume that the user will correct errors at the end of a dictation segment (for example a paragraph). Single words (or in ViaVoice, phrases) can be corrected by double clicking – the user's utterance is played back and a list of suggestions is presented, including the option to delete the word, the user either chooses an option from the list or types in the correct word. The Dragon correction system is less intuitive to the new user and they are encouraged to follow the on-line tutorial which takes them through the process of dictating and correcting text.

5.1.2 Results

The table below gives the recognition rates and text creation speed achieved for all users at their first and best attempt of the three. In general the best attempt was the third one though in some cases there was no difference between the second and third attempts or the second was a percentage point higher. The recognition rate is a simple score of the number of words out of the possible 100 which were recognised correctly first time The final column shows the user's preferred system.

User	Solo 1st attempt	 best	Simply 1st attempt	Speaking best	ViaVoice 1st attempt	 best	Preference
1*	†73% 11wpm	92% 20 wpm	95% 33 wpm	97% 50 wpm			Simply Speaking
2	†66% 20 wpm	98% 33 wpm	91% 25 wpm	98% 33 wpm			Solo
3	76% 11 wpm	96% 25 wpm	†91% 25 wpm	97% 50 wpm			Solo
4	84% 20 wpm	97% 33 wpm	†65% 14 wpm	80% 25 wpm			Solo
5*	†82% 33 wpm	90% 55 wpm			81% 33 wpm	95% 100 wpm	ViaVoice
6*	†81% 50 wpm	95% 67 wpm			82% 33 wpm	82% 67 wpm	Solo
7	71% 11 wpm	96% 44 wpm			†80% 35 wpm	96% 100 wpm	ViaVoice

* had used a speech recognition system before
† used this system first

As can be seen, most users experienced very good recognition rates on each system at their best attempt. There does not appear to be any significance attached to which system was used first. The text creation rates, which include the time taken for error recovery, are also quite acceptable, particularly as this was a totally new experience for four of the users. Perhaps not surprisingly higher speeds were obtainable on the continuous speech system as it is possible to speak faster to it. These rates compare favourably to the time taken by two of the users who were not touch typists to type in and correct the test – 40 wpm and 50 wpm. Users found the set-up instructions fairly easy to follow, although there were some irritations, for example giving the wrong program name in the manual, and giving instructions for putting on the microphone only after you have used it. The time taken to complete the tests was significantly longer for DragonDictate (between 20 and 45 minutes compared to 5 to 15 minutes for Simply Speaking and 10 to 20 minutes for Via Voice) but this is because it is necessary to complete the Dragon tutorial to learn about error recovery which can if required be carried out without recourse to the keyboard. None of the users thought that the set-up time was excessive. All except one of the users (user 4) were very impressed with the performance of the systems and five of them would be prepared to carry out prolonged trials of one of them. The preference expressed was generally for the system on which the user had achieved the best recognition rate and/or the fastest speed. User 3 is an exception, she preferred Solo because she liked to correct as she went along, whilst conceding that SimplySpeaking might be better if you were reading the text. User 6 who had some prior experience of discrete word systems felt that he would need more time to adjust to continuous speech dictation and said that he felt more in control with DragonDictate.

6 Conclusion

The performance of all three systems was impressive with no echo of the traumatic experience of our experiment in 1994 when the new user (No. 1 in our sample) just *could not* get the system to recognise him. We would expect speed and recognition rates to improve over time, though whether most users will be able to obtain the claimed rates of 80wpm for DragonDictate Solo and 140-160 wpm for the continuous speech systems remains to be seen. If we had to choose just one system to meet our requirements it would be DragonDictate Solo – for its cheapness, versatility, and ability to run on a low specification machine. The speed performance of ViaVoice was impressive and it outclasses its predecessor SimplySpeaking, so for straightforward dictation tasks with keyboard corrections this would be a good choice, but only for machines with 32MB of RAM or more. Users who cannot use a keyboard or whose tasks include more than just word processing should choose DragonDictate Solo or possibly the Kurzweil VoicePro although this is only available in the American English version. Both Dragon and IBM also offer more sophisticated versions of their products which would be worth investigating. The most comprehensive currently appears to be Dragon's NaturallySpeaking DELUXE Edition which incorporates all the features of DragonDicate with NaturallySpeaking, providing a totally hands-free environment; this retails at £459, but with the current levels of competition it is likely to get cheaper

7 References

[1] DAMPER, R., and WOOD, S., Speech versus keying in command and control applications, International Journal of Human-Computer Studies, Vol. 42, pp. 289-305, 1995.
[2] MOLNÁR, K., and KLETKE, M., The impacts on user performance and satisfaction of a voice-based front-end interface for a standard software tool. International Journal of Human Computer Studies, Vol. 45, pp. 287-303, 1996
[3] MAIDEN, N.A., NCUBE, C., Acquiring Requirements for Commercial Off-the-Shelf Package Selection, IEEE Software, March 1998.
[4] http://www.endeavour.co.uk
[5] http://www.dragonsys.com
[6] http://www.software.ibm.com/is/voicetype
[7] http://www.lhs.com
[8] http://www.alphasoftware.com
[9] http://www.kurzweil.com
[10] KAY, R., "Do You Hear What I Say?", Byte, January 1998 pp 115-116.
[11] FRENCH, I., HALFORD, P., HEWITT, J., SAPSFORD-FRANCIS, J., Introducing Voice Control – Widening the Perspective, in: Computers for Handicapped Persons - Lecture Notes in Computer Science 860, Zagler, W, Busby, G, and Wagner, R, (eds.), Springer Verlag pp305-312, Vienna 1994.

Session X

Communication

Remote Assistive Interpersonal Communication Exploiting Component Based Development

C.Viglas[1], C.Stamatis[1] and G.Kouroupetroglou[1]

Abstract

Assistive technology has been committed to offer a great deal towards facilitating the communication needs of the disabled. Nevertheless, existing developments tend to focus primarily on face-to-face communication, thus leaving remote communication possibilities unattended or just "put aside". Remote communication for the disabled however, namely the ability to correspond over a public or computer network and especially the Internet, either by transmitting written messages, voice mailing or even participating in a teleconference, is of paramount importance nowadays. In this paper we present the development of a new generation of flexible, open, adaptable, configurable and cost-effective aids supporting both interactive and non interactive remote/ distance interpersonal communication. Details on the requirements, properties, components and services for remote interpersonal communication along with specific component development are given. The implementations presented conform to the recently introduced ATIC (Access to Interpersonal Communication) application development framework which exploits Component Based Software Development technology to maximize modularity and reusability.

1. Introduction

Interpersonal communication, which is taken for granted in our everyday lives, constitutes a crucial issue for many disabled people. Commercially available, computer-based communication aids are usually oriented towards providing partial solutions to specific communication problems for individual users. In most cases they offer from little to none functional flexibility with respect to their adaptation to either different or evolving user needs [2]. Furthermore, current trends demonstrated a prominent preference to matching the disabled user's requirements for face-to-face communication, leaving almost out of reach any consideration for communicating remotely. Namely, the communication aid is considered a computer-based application that solely satisfies the needs of the disabled for communication and interaction with their social or work environment. This in turn, prevents from taking advantage of what contemporary technology offers, namely exploiting the Internet and its services, and thus being able to communicate without barriers with anyone, regardless of the communication partners' physical location or concurrent on-line presence. In other words, harnessing the benefits of electronic mail, messaging or even voice mailing and real-time conversation via tele- or video- conferencing facilities.

It is clear however, that over the past few years the communication requirements of humans have undergone quite a change [8]. Face-to-face communication, although remains highly desirable, is no

[1] University of Athens, department of Informatics, Panepistimioupolis, Ilisia, Athens, GR-15784, Greece,
E-mail: koupe@di.uoa.gr

longer the only predominant need to be matched. People tend to use computer networks more and more utilizing electronic mail and conferencing facilities to *communicate remotely (i.e. from a distance)* with friends and associates from around the globe. In addition, the World Wide Web (WWW) is rapidly becoming the universal delivery medium affecting tremendously the domain of interpersonal communication as well [16]. Furthermore, the developments in electronics and mobile telephony makes reaching and utilizing the telephone device possible for a lot of people with motor disabilities. The contemporary communication aids' market however, has not efficiently tackled these new developments yet.

In this paper we present the development of aids supporting both interactive and non-interactive remote/distance interpersonal communication. The implementations presented conform to the recently introduced ATIC (Access to Interpersonal Communication) application development framework [6] which exploits Component Based Development technology. We have shown previously that the ATIC framework, leads to a considerable change in the way communication aids are looked upon and moreover affects the overall process of developing communicators into more effective products [12].

Our presentation will begin with an introduction of the requirements for remote communication in the assistive technology domain, accompanied by a short presentation of the ATIC framework. The remaining sections will deal with the technical solution we have devised for the implementation of communication aids with remote interpersonal communication support, with emphasis put on describing reusable components.

2. Requirements for Remote Communication for the Disabled

Aiming to develop communicators that would offer additional features than the average contemporary aid, and would be more efficient in dealing with their users' characteristics, thorough studies were conducted, and a number of critical issues were revealed as a result [1]. In particular, the studies focused on the technical specifications of user needs and requirements as well as the features offered by commercially available communication aids. Further analysis of their results led to the identification of the properties for remote interpersonal communication for the disabled that can be mapped into appropriate software components and services.

2.1 Identified Properties for Remote Communication

The communication aid should cater for different modes of communication, including:

- Message exchange (be that voice mail or e-mail, written or spoken) between a disabled person using a symbolic or orthographic language and an able-bodied individual and vice-versa.
- Message exchange between disabled people using different symbolic or orthographic languages.
- Interactive remote conversation (in the form of interactive chat or teleconferencing), between two disabled people or a disable and an able-bodied one, using, again, different languages.
- Use of the communication aid by a speech impaired person to communicate through a telephone device with an able bodied individual using the speech synthesizer of the communication aid in order to output voice to the common telephony network.
- Face-to-face communication between disabled persons and between a disabled person and an able bodied individual.

To design a product that targets on the international market, dealing with not converging user needs in various counties, with not the same cultural backgrounds and using different languages, first of all multilinguality support of a number of alternative/symbolic and natural languages is imperative. As far as adaptability is concerned, the selection of the proper language or communication system and associated user specific vocabulary are crucial. At the same time one has to ensure that people using different languages and systems will not be prevented from communicating (locally or remotely). To maintain effective communication, machine translation should be supported, either word by word or based on a common interlingua meaning, using proper techniques to overcome mismatching vocabularies [3,5]. Multiple output modalities, such as speech, audio, visual, printed, should be supported. The aid should also be adaptable and flexibly configurable at a lexical user interface level [6,15,17], to match the user's motor, sensory and mental needs and abilities. Moreover, it should also offer extensive configuration aspects in terms of: input and output language, input and output devices, and user vocabulary selection, along with a high degree of flexibility and adaptability to the -changing with time- user needs. Needless to say that there are even more issues to be considered, like maintainability and ease of operation.

The aforementioned analysis shows a great diversity as far as disabled user communication needs are concerned, making it extremely unlikely to adopt a general-purpose solution towards communicators' implementation. To achieve this complicated goal, *component software development* is adopted, as it does offer maybe the only viable solution to cost-effective products. This approach offers adequate interoperability to come up with a sufficient range of items/parts ready to be used off-the-shelf, enabling the desired versatility in a communication aid.

2.2 Identified components and services

Further analysis of the discussion presented in the previous section, leads to the following components and facilities that communication aids should support: establish connection over a PTSN network, establish connection over a computer network (namely maintain a TCP Internet connection), handle a communication partner address book, maintain and manipulate a message/mail/voice-mail box, and transmit/receive a message over a network. Such facilities call for either voice output (if connected to the common telephony devices), or proper output modality (text, voice, graphics, etc.) according to the given needs and circumstances if an Internet-based communication is desired.

On the other hand, a number of general purpose components have also been identified and should be considered as they offer basic functionality one would expect to find in any computer-based communication aid. Naturally, the specific components that would comprise a certain communication aid will depend upon the special features and functional characteristics we would like to be present in the communication aid in question. Typical components would support -at least- the following services [3,10]: vocabulary definition, symbol selection, message reception, sentence/message composition, multilingual support and translation between different user languages in a variety of output modalities (text, sound, voice, graphics), message output in a number of ways (such as transmission over a communication channel, speech production, on screen display, and printer output).

Additionally, there exists a set of equally important components also considered as desirable but not always mandatory, providing input acceleration facilities (usually implemented using word prediction techniques), an editor for the construction of symbols of non orthographic (symbolic) languages, and training on the user's communication system or language.

3. Development Framework for Communication Aids

3.1. Component Based Development

Component-based development (CBD) is at present the most widespread technology for application development. It refers to the practice of reaching software solutions by building or buying interoperable components [16]. CBD, is quickly becoming the dominant model for software development, as it has already been ranked most favorable amongst software developers. Supporters exist not only in the generic application domain, but also in the domain of commercial applications and databases. CBD, owes in part its widespread acceptance to the *Internet* which is responsible for almost all upheavals in recent developments. The Internet has influenced commercial and business logic and to support that in a cost-effective manner, component technology seems the ideal solution [16].

CBD, promises and delivers considerable returns of investment in terms of ease of use and training, to both commercial and corporate developers and vendors such as Microsoft, Netscape Communications and Sun Microsystems who are already actively involved in the CBD domain. On the hand, CBD has broken into rehabilitation technology as well, in a quite remarkable manner, having various developers putting a considerable amount of effort particularly in the interpersonal communication domain [4,10,13].

Furthermore, by demonstrating a high degree of reusability (both at the code and application level) and harnessing the advantages of Object-Orientation, CBD has been identified as offering a balanced flexibility and adaptability. Hence, CBD leads to modular and open products that comply to such requirements as: interoperability, modularity, distribution, extensibility, and independence from programming language, as far as building a component application is concerned [9,16].

3.2. Fundamentals of the ATIC approach

The implemented ATIC framework, aims to be used by a number of different vendors, harnesses the benefits offered by the component-based technology, and introduces a new perspective in building better, feature-rich and more manageable communication aids, affecting drastically their life cycle [11,12]. It endorses a modular design for communication aids complying to a general adaptable solution to serve a wide range of users, able to deal adequately with users having motor, perceptual and cognitive problems. A versatile technology is used [11], allowing for rapid changes, and supporting multiple features in order to meet the physical abilities, cognitive and language levels and conversational needs of persons using these aids [17]. Any user language, symbolic or natural, is supported via a database of multimedia language elements, to be utilized in all aspects of the disabled user's communication needs, be that face-to-face or remote communication and even group conversations (i.e. support for multiple output modalities). In order to reach the aforementioned goals, the ATIC approach utilizes and is based upon, the provision of a novel component architecture (defining communication protocols, components connectivity, building rules, and software tools inclusive) [10,12], and the implementation of a number of architecture-compliant components (modules), that can be interoperably used to built communication aids.

All in all, ATIC constitutes a development framework, which promotes reusability in the design, the code generation and binary execution. These features render our framework in an advantageous position as far as communication aids application development is concerned. Taking advantage of its Object-Oriented design [11] and the features of CBD, the ATIC framework incorporates a set of notable advantages that are often demonstrated by object-oriented frameworks [7] leading to:

- better and more concrete control over the application development process,
- greater reusability and adaptability potential for the developed components,
- applications that can demonstrate either a general purpose or a highly specialized character,
- even more robust and customizable user interfaces,
- ultimately, bigger productivity and variety of implementation, which can lead, in the long run, to a vaster range of choice but also to better quality (due to competition), which is of utmost importance as far as interpersonal communication for the disabled is concerned.

According to ATIC, any identified user needs are mapped to appropriate communication functions and translated into the software-oriented domain. The main functions identified are then broken down into sub-functions leading to *elementary functional elements* (i.e. *services*) that the communicator's components would be expected to offer and serve]. In this context, a *communicator is considered a system providing a number of functions and/or services dependent on the particular user-needs, abilities and cognitive level. Each function may be implemented independent of others, either as a separate entity or as a set of elementary services, offered by a component or set of interoperable components responsible for the implementation of a function or service in a way transparent to the architecture and the communicator itself.*

The ATIC framework defines both the programming model and the required binary standard for creating, managing, and accessing object-based components that provide (and use) services to other objects and applications. As a technology, it allows components to interact across process boundaries as easily as objects interact within the same process. This is enabled by means of a message registration and passing mechanism. ATIC does specify an object model that supports concurrency, re-entrant multi-threading, internal synchronization of processes and inter-process communication of components, which allow for component applications to be built using virtually any programming language [11].

4 Remote Interpersonal Communication Implementation

Remote communication was successfully implemented under the ATIC component software framework, and its *Object Model (using OMT notation [14]) being presented in Figure 1*. Two different kinds of objects can be distinguished: the *module* object, which represents a component of the interpersonal communication aid responsible for the implementation of a specific function, and the *ATIC* object. A more general object, the *Communicator*, represents the entire communication aid.

The *ATIC* component is responsible for any information exchange, plus error handling, between the communicating components. All information between the components and the *ATIC* is implemented through message exchange: *ATIC* receives messages from the "client" component, decides which component is the proper recipient (more than one module might be affected by a certain message), and transmit the necessary message to the "server" module(s). This communication is represented through the "message exchange" association [11].

Aggregation is identified between the *Communicator* object the *ATIC* and the component objects. More specifically, the *Communicator* is an abstract object, which consists of one ATIC instance and a collection of instances of the Module class. For this reason, the various instances of the Module object, which are most crucial in the model, are discussed in the following sections, using top-down generalization. Subclasses of the Module class can be derived by, identifying the desired functionalities of a most generic (all-encompassing) communicator, and then decomposing them into elementary functions. Each function is independent of the others and unaware of their presence or absence in the communicator system (if a service that is not provided *is* required, the relevant *ATIC* component is responsible for sending the appropriate error messages). This is the reason why no association or link whatsoever exists between *Module* subclasses. Communication between them is established through message exchange via the *ATIC component*.

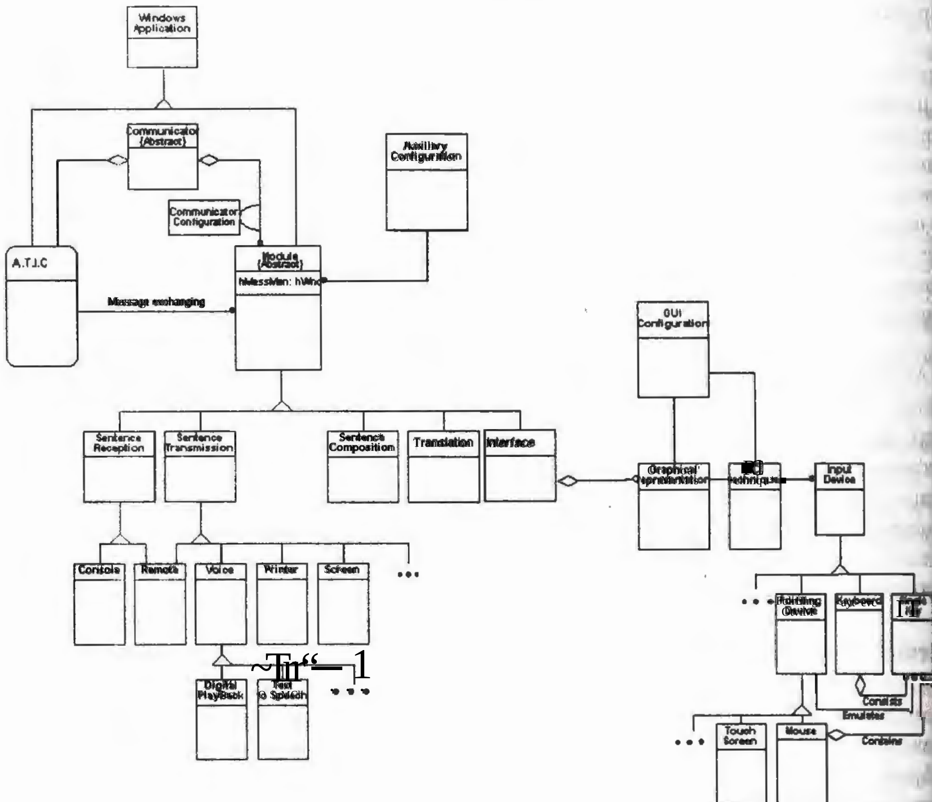


Figure 1: The Object Model for remote communication

4.1 The Module object

The Module object is generally concerned with the implementation of certain services that the communicator is expected to provide. Its functionality and specifications depend upon the particular component that is being developed.

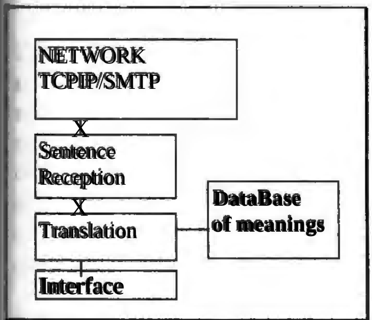


Figure 2: Message reception

Considering remote communication, two different functional modes can be identified; *the sentence reception mode* and *the sentence transmission mode*. The first is accomplished through the underlying network by using the SMTP protocol *in the case of E-mail (Figure 2)*. The message is received from the SMTP port by the Sentence reception component. This component firstly removes the header of the message and then it breaks the message in a set of discrete meanings associated with the original message composed in the user's own language. The meanings are passed through the ATIC component to the translation component by the use of a message named TRANSLATE. The translation component, incorporating a database of meanings, translates the meaning to the proper user language and passes it to the interface where the message appears in an editor. Almost the same scenario takes place in the case of interactive chat. The difference is that the received message is broken down in meanings, each individually carried in TCP/IP Internet packets. Then, each meaning is passed to the translation component and presented in the proper user language. This process takes place in real time as the meanings composing the user message arrive.

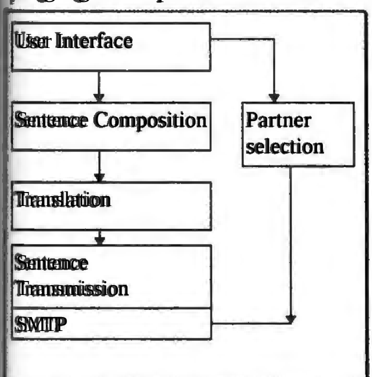


Figure 3: Message transmission for e-mail

Sentence transmission mode is more complex. It begins with the user taking advantage of the services of the message composition component through the user interface component. The message is broken in different words. The discrete words are passed through the ATIC component to the translation component by use of the TRANSLATE message. The translation component using the database of meanings translates the message to a set of meanings which are passed to the proper network protocol, such as the Simple Mail Transfer Protocol (SMTP) in the case of e-mail and TCP/IP in the case of interactive chat. Note that the sentence transmission component is different in each of those cases. *In the first case (Figure 3)*, the sentence transmission component passes the whole message (as one chunk of text) to the mail protocol. The selection of the user's communication partner is handled by the User Interface module, where the user chooses a graphical representation of his/her partner. This is then mapped to the corresponding e-mail address by the Partner Selection component and then passed to the mail protocol. In the *second case (Figure 4)*, the message is passed word-by-word to the TCP/IP layer which in turn acknowledges reception to the Sentence Transmission component. In the case of a communication error the, User Interface component gets a notification of the error that occurred. Selecting a communication partner in this case is also handled by the User Interface where the user chooses a graphical representation of his/her partner. This is then mapped to an IP address by the Partner Selection component, which is responsible for passing it to the Sentence Transmission component. Then this IP

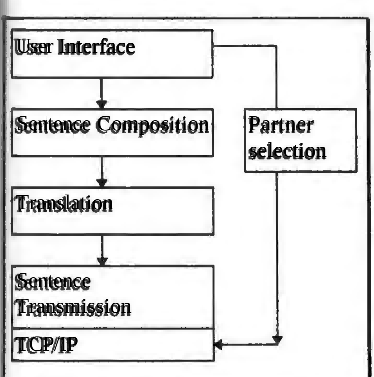


Figure 4: Message transmission for chat

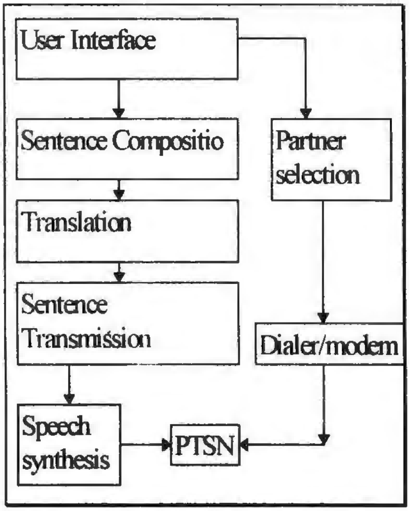


Figure 5: Message transmission for PTSN connection

address is used by the Sentence Transmission component to establish connection, send, listen, check and retransmit packets utilizing the TCP/IP layer services. *In the third case scenario (Figure 5)*, the Sentence Transmission component passes the composed message to a speech synthesizer which in turn produces output to a telephone device. The communication partner selection is once more achieved in the same manner as in the previous cases. The Partner Selection component utilized in this case invokes a modem device to dial directly the desired partner. As soon as the connection is established the User Interface component is notified so that the user can start composing his/her message.

4.2 The User Interface Component

One of the most important components of the communication aid is the User Interface component. This component is compliant with the ATIC component software architecture. This means that it communicates with the other components using ATIC messages. The User Interface component registers the INTERFACE message, which carries a structure containing the specific command issued by a functional component to be handled and intended for the Interface object. One of the biggest obstacles that have to be overcome in order to provide a generic solution to the user interface design in a component based environment, is managing the GUI of the application. This becomes more evident in a communication aid where the user interface has to have special behavior supporting many different scanning techniques and interaction modes. Taking into account that the user interface objects must respond to special devices and that different components will have to interact as far as user interface is concerned, having each functional component developer to embody the user interface into each component would make building a communicator extremely complex. The solution to that is to provide a separate User Interface component with a high level GUI definition language that would be interpreted in real time by a user interface server in order to provide and manage the GUI.

```
Screen 1{
backcolor R,G,B
group1{
button1,x,y,z,w,
R,G,B.bitmiap.bmp,ATICMESSAGE,\wparam,Lparain
.
.
butromi,x,y,z,w,
R,G.B,bilniap.bmp,ATICMESSAGE,wparam,Lparam
List1. x,y,z,\w
.
.
Edit1x,y,z,w

Group1.1 {
Group 1.1.1
Group {1.1....1i
)
}}
{
group2(
}}
.
.
groupn{
}}
}}
Scrren2{
}}
.
.
Screen n{
}
```

Figure 6: Example of the user interface definition language

The User Interface component of a communication aid is structured, namely, there is a hierarchy amongst the interface objects that comprise it. Thus, the interface definition language should be a structured one supporting multilevel encapsulation of objects. The hierarchy defined in this encapsulation determines the order in which the interface objects are scanned. The first level of this encapsulation is the "screen" object. The Screen is the *container*, which contains all the other objects. Within a screen, scanning groups are defined. The scanning groups are objects that can contain other groups or childless objects. A Childless object is a button or a listbox or an editor. *An example of the GUI definition language is presented in Figure 6.*

5. Discussion

The remote interpersonal communication aids described above have been implemented under the MS-Windows 95 environment for an IBM-PC compatible platform. They offer enhanced functionality by fulfilling all the requirements identified in section 2 of this paper.

The communication aids described in this paper were evaluated by speech-motor and language-cognitive impaired users in Great Britain and Finland within the ACCESS project as well as in Greece under the HORIZON project HESTIA.

Figure 7 depicts a typical remote communication scenario for two disabled people using ATIC compliant communication aids. More Specifically, they reside different countries (Greece and Finland in the example), one is using BLISS as a communication system while the other is using Pictogram, and they establish a connection over a telephone or computer network.

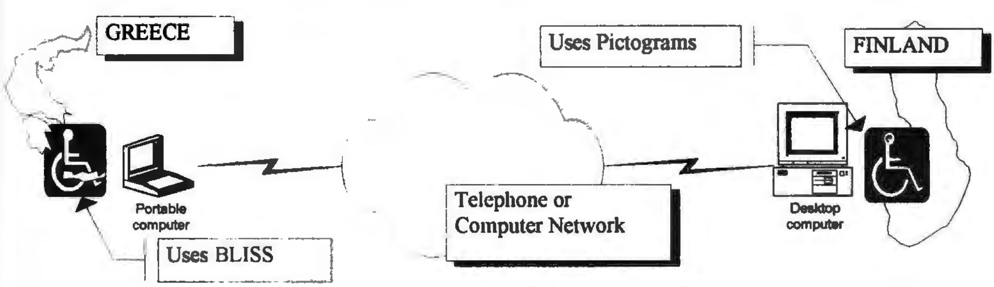


Figure 7: Typical ATIC enabled, remote communication

Acknowledgments

The work reported in this paper, was carried out within the framework of the ACCESS project, partially funded by the TELEMATICS for the Integration of Disabled and Elderly people (TIDE) Programme of the Commission of the European Union, DG XIII. Partners in this consortium were: CNR (Italy), FORTH (Greece), University of Athens (Greece), RNIB (U.K.). Seleco S.P.A. (Italy), MA Systems and Control (U.K.), Hereward College (U.K.), NAWH (Finland), VTT (Finland), Pikomed (Finland) and University of Hertfordshire (U.K.).

References

[1] AHONEN, J., A.Kalliala, E.Kaasinen, G.Kouroupetroglou, R.Lauder, K.Nyholm, T.Pulli, V-P.Raty, A-L.Salminen, K.Salomaa, C.Stephanidis: "*Final Report on User Needs*", TIDE project 1001 ACCESS, Report D.4.1.2, Volume II. February 1995.

[2] ALM, N., A.F. Newell, J.L. Arnott. An Integrated Research Strategy for Improving Communication Systems

for Severely Physically Impaired non-speaking People. Proceedings of the 1st TIDE Congress, pp. 249-253, 1993.

[3] ANTONA, M., C. Stephanidis, and G. Kouroupetroglou, Vocabulary Management in Modular Interpersonal Communication Aids, in: Advancement of Assistive Technology (AAATE) G.Anogianakis, C.Buhler and M.Soede (eds.), IOS Press, 1997, pp 200-205.

[4] DEMASCO, P., E.J. Ball, S.Tywand, D.Blodgett, W.Bradley, J. Dunaway, Z.Kazi, Towards Modular AAC Software: An Object-Oriented Architecture, RESNA International '92, June 6-11, 1992, pp. 119-121.

[5] DORR, B.J. Machine Translation: A View from the Lexicon, MIT Press, 1993.

[6] EMÍLIÁNI P.L., J.Ekberg, G.Kouroupetroglou, H.Petrie, and C.Stephanidis, The ACCESS Project: Development Platform for Unified Access to Enabling Environments, Proceedings of ICCHP' 96, (International Conference on Computers and Handicapped People), Linz Austria, October 1996, pp. 69-75.

[7] FAYAD, M.E., and D.C. Schmidt. Object-Oriented Application Frameworks, Communications of the ACM, Vol. 40, No. 10, October 1997, pp. 32-38.

[8] GOODENAUGH-TERPANTIER. Design Goals for Augmentative Communication, Assistive Technology, Vol. 6.1/1994, pp. 3-9.

[9] JOHNSON, R.E., and B.Foote. Designing reusable classes. J. Object-Oriented Programming 1, 5 (June/July 1988), pp. 22-35.

[10] KOUROUPETROGLOU G., A.Paramythis, A.Koumpis, C.Viglas, A.Anagnostopoulos and H.Frangouli, Design of Interpersonal Communication Systems Based on a Unified User Interface Platform and a Modular Architecture, Proceedings of the TIDE Workshop on User Interface Design for Communication Systems, Brussels, July 7, 1995, pp. 8-17.

[11] KOUROUPETROGLOU G., C.Viglas, A.Anagnostopoulos, C.Stamatis and F.Pentaris, A Novel Software Architecture for Computer-based Interpersonal Communication Aids, Proceedings of the 5th ICCHP'96 (International Conference on Computers and Handicapped People), Linz Austria, October 1996, pp 715-720.

[12] KOUROUPETROGLOU G., C.Viglas, C.Stamatis, and F.Pentaris, Towards the Next Generation of Computer-based Interpersonal Communication Aids, in: Advancement of Assistive Technology, G.Anogianakis, C.Buhler and M.Soede (eds.). IOS Press, 1997, pp 110-114.

[13] LUMDALV, M., D. Svanaes, Comspec - Towards a Modular Software Architecture and Protocol for AAC Devices, 1st TIDE Congress, 1993.

[14] RUMBAUGH, J., M.Blaha, W.Premerlani, F.Eddy and W.Lorensen, Object-oriented modeling and design, Prentice Hall, 1991.

[15] SAVIDIS, A., A. Paramythis, D. Akoumianakis, C. Stephanidis, Designing user-adapted interfaces: the unified design method for transformable interactions. ACM Conference on Designing Interactive System: Processes, Methods and Techniques (DIS '97), Amsterdam, The Netherlands, 18-20 August 1997, pp. 323-334.

[16] SPITZER, T., Component Architectures, DBMS, Vol. 10, No. 10, September 1997, pp. 56-68.

[17] STEPHANIDIS, C., and G.Kouroupetroglou, Human Machine Interface Technology and Interpersonal Communication Aids, Proceedings of the Fifth COST 219 Conference, Tregastel-France, June 7-8, 1994, pp. 165-172, ISBN 951-33-0004-8.

[18] STEPHANIDIS, C., A. Savidis, D. Akoumianakis, Unified Interface Development: A step towards User Interfaces for All, in: Advancement of Assistive Technology, G.Anogianakis, C.Buhler and M.Soede (eds.), IOS Press, 1997, pp. 29-33.

International

One Basic Research Approach for Assistive Technology in Japan

Tohru Ifukube,
Shuichi Ino, Norihiro Uemi,
and Chikamune Wada[1)]

Abstract

One basic research approach for assistive technology in Japan is proposed in this paper. Ifukube (one of the authors) has developed this research approach through about 25 years' experience in assistive technology studies. He and his co-researchers have designed several aid devices as reviewed in this paper and have obtained many basic findings concerning human information processing. Some of the devices have been manufactured in Japan and the technologies as well as the basic findings have been applied to construct both virtual reality systems and robots. Moreover, newly developed virtually reality and robotics technologies have been fed back to design better models for developing assistive devices.

1. Introduction

We have been carrying out the assistive technology studies, especially communication aid studies, for about 25 years, and have designed several aid devices which are in practical use or will be put into use. Moreover, we have obtained many findings concerning sensory integration, concept formation and sensory-motor association in the human brain. We will refer to our research regarding the aid devices and mention how that research has been related to virtual reality and robotics. Finally, we will introduce our research approach to developing assistive technology.

2 Assistive devices designed by authors

2.1 Tactile voice coder

Through fundamental research on auditory and tactile information processing, we developed a tactile voice coder for the deaf about 21 years ago which has been manufactured in Japan. Our research was broadcast as a documentary program called "my finger can hear letters". The tactile voice coder is a device which produces sound spectral patterns that are analyzed in 16 frequency components. When an index fingertip touches a piezo-electric vibrator array (consisting of 16 rows by three columns as shown in figure 1), the device makes it possible to discriminate the

---

1) Laboratory of Sensory Information Engineering, Research Institute for Electronic Science, Hokkaido University, Sapporo 060, JAPAN

first and the second formants of vowels as well as some consonants such as fricatives. It assists in lip reading and also helps to obtain feedback for speech training. It has already been put to practical use in Japan, and studies on its evaluation are being at some institutes for the deaf. From identification tests of Japanese monosyllables consisting of /ku/, /su/, /nu/, /fu/, /ru/, /tsu/ and /yu/, the identification rate increased from 23% (lip-reading only) to 68% (lip-reading plus tactile voice coder) after one week of training [2]. In cooperation with an electronic company, we are now improving the tactile voice coder using a digital signal processor (DSP) to make it smaller and cheaper [18].

The findings and technologies regarding tactile voice coders will also be applied to a tactile display for virtual reality systems.

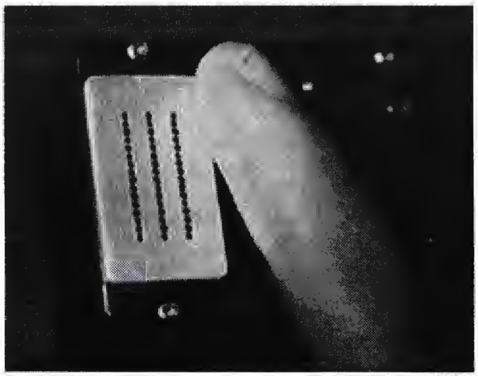


Fig.1 *A vibrator array for tactile voce coder*

2.2. Voice Typewriter

About 16 years ago, the tactile voice coder was applied to a voice typewriter for acquired deaf people for whom it is difficult to learn lip-reading and finger language. This device can convert monosyllabic voice sounds such as /a//ka//sa//ta//na//ha/ into Japanese letters almost in real time and shows them on a display. There are only 5 vowels and 14 consonants in Japanese. Every Japanese word is pronounced as a series of monosyllables. For example, Tokyo is pronounced as /to//u//ki//yo/ and /u/. Every monosyllable corresponds to one Japanese letter. This makes it easy to design a voice typewriter.

Fig.2 *Voice typewriter*

Figure 2 shows the voice typewriter in which a microprocessor was used. Some algorithms were performed by the hardware to decrease the conversion time. Its recognition rate was 96% and its response time was about 0.2 [3]. The voice typewriter is a substitute for finger language or conversation writing. It can also, be used as a typewriter for the upper-limb disabled. The voice typewriter has been applied to the input device of a Japanese word processor. We are now improving the voice typewriter using a micro-processor to make it smaller and cheaper.

2.3. Cochlear Implant

We have been studying auditory prostheses which stimulate surviving auditory nerves of the deaf who have lost the function of hair cell receptors inside the cochlea. However, it was almost impossible to continue the study in Japan because of a lack of co-researchers. Therefore, one of the authors of the study (prof. Ifukube) went to Stanford University to continue the study of auditory prostheses about 13 years ago. In our auditory prosthesis, auditory nerve cells were electrically stimulated by 8 electrodes placed in the cochlea as shown in figure 3.

Spectral patterns were then analyzed in 8 steps. The 8 signals were transmitted electro-magnetically from a transmitter outside the body to a receiver inside the body [4][5]. However, when a tester's voice stimulated the auditory nerves of a female patient, she reported that the perceived signals sounded very different from voices which she remembered from the time she had normal hearing. Furthermore, their ears were infected twice during prof. Ifukube's stay at Stanford.

Fig.3 *Implanted electrode array*

Although our method was not so effective for recognition of speech, especially consonants, we have acquired some other findings regarding effectiveness of auditory electrical stimulation [14][6].

2.4. Tinnitus Suppresser

Fig.4 *Implantable tinnitus suppresser*

It has been known that fifty or sixty percent of hearing impaired patients suffer from tinnitus which causes them stress and insomnia. The fact that electrical stimulation to the cochlea suppresses tinnitus to some extent was observed in the course of developing our cochlear prosthesis for the deaf [16]. We have developed a tinnitus suppresser which we have used to treat tinnitus patients. With this treatment, about 30 % of the patients claimed that the tinnitus had disappeared without any side effects. However, since this method suppresses tinnitus for only a short period of time, many tinnitus suffers have long been waiting for an implant system which can suppress tinnitus whenever it appears.

The implantable tinnitus suppresser which we developed consists of a wave generator, a primary coil outside the ear, a secondary chip coil implanted under the skin of an external ear and a Pt-Ir electrode placed on the promontory of the cochlea as shown in figure 4. The secondary coil connected with an electrode was implanted for 7 patients [13].

2.5. Digital Hearing Aid

In general, elderly people who have suffered from a hearing impairment have less ability to understand spoken language even though they can hear the speech sounds. This phenomenon seems to be due to a decrease in the recognition of auditory time patterns in the speech area of the cortex. Therefore, in cooperation with the Central Research Laboratory of the Hitachi company, we have designed a hearing aid which can slow down speech without any pitch frequency change by using a digital signal processor as shown in figure 5 [15].

Fig. 5 *Digital hearing aid*

This hearing aid was evaluated and improved by our university. In our hearing aid, first, the pitch frequency of a vowel part is extracted; then the pitch wave is repeated so that the duration of the vowel part can be slowed down. We have proved that this device is effective in catching the meaning of rapidly spoken sentences for the elderly sensorineural hearing-impaired. This hearing aid has been manufactured by Hitachi LTD.

2.6 Electric Artificial Larynx

Various methods for vocal rehabilitation have been applied to patients who have undergone laryngectomee, thereby losing their speech function. We have studied the vocalization mechanism of a mynah bird, which can imitate the human voice, in order to apply it to a synthetic sound generator for such patients [7]. An electric artificial larynx is one of the artificial larynges for the people who have failed to master other forms of substitute speech. However, this prosthesis makes it difficult to produce natural voices. From the analysis of the mynah's vocalization mechanism, we have found that the mynah can clearly imitate phonetic information such as intonation and accent. This is the reason why mynahs' voices can be heard as natural voices by humans. In order to improve a conventional electric artificial larynx, we have proposed a new method that can allow laryngectomee patients to control intonation by using their respiration [12].

The device consists of three parts as shown in figure 6. The first part is a pressure sensor that can detect exhaled air pressure produced from a stoma made by a surgical incision into the neck. The second part is a transformation circuit made of a micro-processor that can convert air pressure into a pitch frequency for voice. The third part is an electromechanical vibrator that can be attached to the neck. By using the optimal parameters, the pitch pattern of the electric larynx voice became clearly similar to the pattern produced from a normal subject after one day of training. This type of electric artificial larynx has just been manufactured by a company in Japan. It has been proven that intonation is very important to make an artificial larynx voice natural [13].

Fig.6 *Electric artificial larynx* company

2.7. Ultrasonic Eyeglasses

A new model of a mobility aid for the blind has been developed using a microprocessor and ultrasonic devices. In this model, a down swept frequency modulated ultrasound is emitted from a transmitter with broad directional characteristics in order to detect obstacles as shown in figure 7.
Ultrasound reflections from the obstacles are picked up by a two-channel receiver. The frequency of the emitted ultrasound is swept from 70 to 40 kHz within 1 ms, giving it almost the same characteristics as the ultrasound which an FM-bat produces for echo-location. The frequency of the reflected ultrasound wave is reduced by about 50 : 1 by using a micro-processor with A/D and D/A converters. These audible waves are then presented binaurally through earphones. In this way, obstacles may be perceived as localized sound images corresponding to the direction and the distance of the obstacles.

Fig.7 *Ultrasonic eyeglasses*

With this device, a blind person can recognize a 2-mm-diameter wire at a distance of about 1 meter. We also found that the blind could discriminate between several obstacles at the same time without any virtual images. This mobility aid, modeled after the bats' echo-location system, is very effective at detecting small obstacles placed in front of the head [8].

However, most blind people can detect obstacles without these devices. This ability is called the "obstacle sense". Figure 8 shows two steps involved in the obstacle sense: first perception (2) and final appraisal (3). We have been investigating the mechanism of the obstacle sense based on psychophysical experiments using blind students. We have found the reason why the blind can detect the obstacle is that they can discriminate the tiny difference of sound field between with obstacle and without obstacle [7]. Furthermore, we were able to make the blind hear "virtual obstacles" by controlling the sound field produced from a speaker array. This study has also been related to virtual reality research. In the near future, we intend to design a mobility aid device which will use the ability of the obstacle sense.

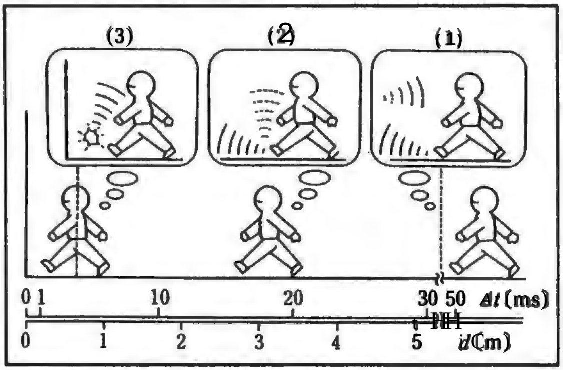


Fig.8 *Two steps of the obstacle sense*

2.8. Transfer Aid

We have designed a transfer aid which can move a bed-ridden patient from a bed to a wheel chair or from a wheel chair to a bed by using a metal hydride actuator. A metal hydride (MH) actuator which uses the reversible reaction between heat energy and mechanical energy of a hydrogen absorbing alloy, has recently attracted much attention. Hydrogen absorbing alloys are capable of storing as much hydrogen gas as approximately 1000 times their own volume. By heating the alloy, hydrogen equilibrium pressure increases and hydrogen is desorbed, whereas by cooling the alloy, hydrogen equilibrium pressure decreases and hydrogen is absorbed. In this


Fig.9 *Transfer aid*

way, it is possible to utilize the mechanical energy of hydrogen gas pressure by manipulating heat.

A Peltier element is used as a heat source, and by changing the direction of the electric current to the element, the alloy is heated or cooled. The functioning part uses metal bellows, which insulate the hydrogen.

The drive function, which uses hydrogen absorption and desorption, has a buffer effect and prevents extreme power changes or shock. Thus, this MH actuator is milder for humans and more suitable for use in equipment which is attached to people. Figure 9 shows a transfer aid in which one MH actuator is used [19][17]. We have also developed a robot hand with tactile sensory feedback as shown in figure 10 [9][11] and a wheelchair with a lifter using an MH actuator with a 40 g alloy.

Fig. 10 *Tele-existence robot hand*

10. Virtual Reality Studies [10]

Based on the above techniques and the findings, we have constructed a virtual reality (VR) system which consists of a head mount display, a speaker array, a rotational chair, and a sound proof room as shown in figure 11. The temperature of the room can be changed in the range from minus 4 degrees to plus 40 degree. We are now investigating mechanisms of sensory integration. As one example, we have investigated how the rotational stimulation influences the visual and auditory sensations. These kinds of findings will be useful to design both better models of VR systems and rehabilitation devices.

Fig. 11 *Virtual reality system*

A research project of a mixed reality (MR) system was begun last year in order to construct a kind of augmented reality in cooperation with the Ministry of International Trade and Industry (MITI), and Canon company. MR involves integrating elements from the real world with a virtual world. Information from reality and virtual reality is put together and displayed. In order to create a smooth MR environment, a number of problems have to be dealt with.

We have been asked to evaluate how the problems influence the human body; In particular, the visual function, the autonomic nervous system, and the sense of equilibrium. In order to investigate these influences, the Sapporo research branch was constructed. In the center, we have three rooms: (1) a control room in which there are some computers such as an Onyx 2 to create images and control equipment, (2) a display room in which an arch-screen (figure 12), a motion base with two force plates , a speaker array and a 3-D motion analysis system have been installed, and (3) a bio-medical measurement room in which we can measure the visual function, the autonomic nerve function and the sense of equilibrium using several kinds of equipment.

Fig.12 *Mixed reality system*

We are planning to investigate how sensory integration is formed in the human brain when different stimulations such as the visual and the balance stimulations are displayed simultaneously or successively. The findings that will be obtained from the research project of MR will be useful in creating better aid devices.

3. Conclusion

It can be said that present-day robots and computers are also "disabled" from the point of view of their sensory-motor functions. The basic findings concerning human sensory-motor functions and aid devices for the disabled will be useful for designing virtual reality systems and robots.

In the near future, the newly developed technologies of virtual reality and robots will be applied to design better models of communication and mobility aid systems for people with sensory-motor disorders. This is our research approach of assistive technology as schematically shown in figure 13.

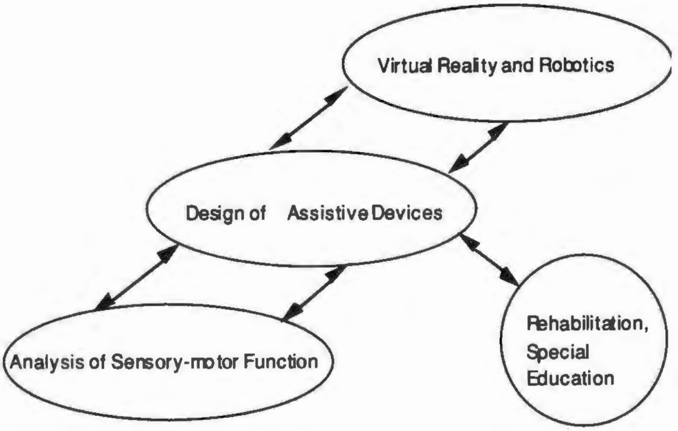


Fig. 13 *One basic research approach for assistive technology*

References

[1] AOKI, N., IFUKUBE., 1/f shimmer and jitter in normal sustained vowels and their application to speech synthesis, in: J.Acoust.Soc.Am (In press) (1998)

[2] IFUKUBE, T., A cued tactile vocoder, in: Computers in Ading the Disabled, J. Raviv (ed), North-Holland Publishing Co., 197-215 (1982)

[3] IFUKUBE, T., Design of a Voice Typewriter, in: T. Ifukube, CQ Publishing Co., (195 pages, in Japanese), (1984)

[4] IFUKUBE. T., Discrimination of synthetic vowels by using tactile vocoder and a comparison to that of an eight-channel cochlear implant, in: IEEE Trans., BME 36(11) 1085-1091 (1989)

[5] IFUKUBE, T., WHITE, R. L., A speech processor with lateral inhibition for an eight channel cochlear implant and its evaluation, in: IEEE Trans., BME 34(11) 876-882 (1987)

[6] IFUKUBE, T., Signal processing for cochlear implants, in: S. Furui & M. M. Sondhi (ed), Advances in Speech Processing, Marcel Dekker, inc., New York, 269-305 (1992)

[7] IFUKUBE, T., Sound-based Assistive Technology for the Disabled, in: T. Ifukube, Corona Publishing Co., (250 pages. in Japanese) (1997)

[8] IFUKUBE, T., SASAKI, T., PENG, C., A blind mobility aid modeled after echolocation of bats, in: IEEE Trans., BME 38(5) 461-465 (1991)

[9] IFUKUBE, T., A study on tele-exsistence of the tactile, in: J. Robotics and Mechatronics, 9(3) 197-202 (1997)

[10] IFUKUBE, T., From sensory substitute technology to virtual reality research, in: Artificial Life and Robotics, Springer-Verlag Publishing Co., (invited paper, 10 pages, in press)(1998)

[11] INO, S., IZUMI, T., IFUKUBE, T., Design of an actuator for tele-existence display of position and force to human hand and elbow, in: J. Robotics and Mechatronics, 4(1) 43-48 (1992)

[12] UEMI, N., IFUKUBE, T., Design of a new electrolarynx having a pitch control function, in: IEEE Inter. Workshop on Robot and Human Communication, 198-203 (1994)

[13] MATSUSHIMA, J., MIYOSHI, S., IFUKUBE, T., Surgical method for implanted tinnitus suppresser, in: Inter. Tinnitus Journal, 2(1) 21-25 (1996)

[14] MIYOSHI, S., IFUKUBE, T., MATSUSHIMA, J., Proposal of a new method for narrowing and moving the stimulated region of cochlear implants -animal experiment and numerical analysis-, in: IEEE Trans., BME (in press) (1998)

[15] NEJIME, Y., ARITSUKA, T., IFUKUBE, T., MATSUSHIMA, J., A portable digital speech-rate converter for bearing impairment, in: IEEE Trans., Rehabilitation Eng., 4(2) 73-83 (1996)

[16] SAKAJIRI, M., IFUKUBE, T., MATSUSHIMA, J., A method for suppressing tinnitus by electrical stimulation to cochlea and remedial value, in: J.Acoust. Soc. Jpn, (English) 14(6) 453-455 (1993)

[17] TANAKA, T., INO, S., IFUKUBE, T., Objective method to determine the contribution of great toes standing balance and preliminary observations of age-related effects, in: IEEE Trans., Rehabilitation Eng., 4(2) 84-90 (1996)

[18] WADA, C., IFUKUBE, T., Proposal and evaluation of the sweeping display of the speech spectrum for a tactile vocoder used by the profoundly hearing impaired, in: Electronics and Communication in Japan, 79(1), 55-66 (1996)

[19] WAKISAKA, Y., MURO, M., INO, S., IFUKUBE, T., Applications of hydrogen absorbing alloys to metal and rehabilitation equipment, in IEEE Trans., Rehabilitation Eng., 5(2) 148-157 (1997)

CHINESE DISABLED PERSON INFORMATION RETRIEVAL SYSTEM AND RELEVANT WORK

Quankai Li

Abstract

Chinese Disabled Person Information Retrieval System contains information about survey of persons with disabilities , disabled persons" organizations , laws and regulations , national programs, outstanding disabled persons, education, cultural life, poverty alleviation, welfare and etc. It is developed by China Rehabilitation Research Center. The establishment of the system is sponsored by China Disabled Persons" Federation and its aim is to assist the government for the development of disabled persons undertakings as well as to serve disabled persons and their organizations.

1. Introduction

There are more than 60 million persons with disabilities in China. They constitute one fifth of the total families in China. China, as a country with the largest disabled population, is fully aware of its responsibility and the role it should play.

China Rehabilitation Research Center (abbreviated CRRC), a state owned institution affiliated with the China Disabled Persons' Federation , serves to collect , analyze and exchange information on disabled persons and on rehabilitation services for the disabled persons so that a national information network for the disabled will be established.

Five years ago, Disability and Rehabilitation Database in Chinese Language (abbreviated DR)), the first database made for disabled persons in China, was developed by CRRC specifically for physicians specializing in the practice of medical rehabilitation . The records we produced in DR database were collected from 400 journals ,100 newspapers, conference papers and laws published in China. In the past several years, DR database has been used by people who are interested in medical rehabilitation in China. Through individuals' searches, DR database is found lack of policies, education , national program, advocacy and legislation and other information relevant to disabled persons . We decided to make the second database for disabled persons in China -- Chinese Disabled Person Information Retrieval System (CDPIRS) , which relates to disabled persons' social problem except medical rehabilitation. This program got financial support from China Disabled Persons' Federation., a national organization for all categories of disabled persons in China, which established in March 1988 with the Chinese government approved and mandated by law . China Disabled Persons' Federation represents and protects the legitimate rights and interests of the disabled as well as provides services to the disabled, and at the same time, assists the government in managing and coordinating the work for the disabled with the functions of representation, service and management. China Disabled Persons' Federation has established local branches at the provincial, municipal, county and township levels.

2. Method

2.1. CDPIRS's operating environment

Available hard disk space 14 MB and IBM386 compatible PC or above

2.2. Data structure

Structuring data design of the database consisted of 2 files: FDT(field definition table),PFT(print format table).Once the FDT was produced, the data structure of CDPIRS was generated. CDPIRS is easier to be used than Disability and Rehabilitation Database.

2.3. Range of selection

Documents, summary , reports , newsletters, stories of outstanding persons , lecture notes and statistical tables with the China Disabled Persons' Federation approved and other important documents relevant to the course of disabled persons.

Formal and informal newspaper , periodical and non-periodical edited by local organizations of Disabled Persons' Federation at the provincial ,municipal, county and township levels.
Books relating to the field of disabled persons.

Important articles about disabled persons published on local newspapers.

Documents collected by Ministry of Civil Affairs , Ministry of Public Health , Ministry of Labor , State Education Commission and Ministry of Justice.

3. Fields of each record

Below is the CDPIRS fields, their corresponding codes and full descriptions of 7 essential fields. 16 fields are definited in the field definition table.

3.1. *010 publication type*

The 010 field indicates the kind of information or the format in which it is presented. Examples of publication types include:

I newsletter
C conference
D official document
T table
P periodical and non-periodical
B book
W writing cut from newspaper
E etc.

3.2. *111 accession number*

The 111 field contains the unique eight-digit number assigned by the Library of China Rehabilitation Research Center(LCRRC).The first two digits represent the province of the published article. The second two digits represent the year of entry.

3.3. ***200 title***

The 200 field contains the title of the article.

3.4. ***250 address of author***

The 250 field contains the address of author.

3.5. ***300 author(s)***

The 300 field contains author(s) of the article.

3.6. ***470 source***

The 470 field contains the complete bibliographic citation, including journal title, date of publication, volume number, issue number.

3.7. ***620 Chinese Subject Headings(CSH)***

The 620 field contains the controlled vocabulary terms or subject headings assigned by indexes at Institute for Chinese Scientific Information and is used to group documents on similar topics . Chinese Subject Headings are standard vocabulary terms that describe the concepts covered each article in the CDPIRS . Chinese Subject headings allow users to retrieve all references to a particular topic, even if different terminology is used in these records. Each record contains several CSH. After each article has been indexed ,it will be put into the CDPIRS .

4. Search

4.1. Basic search and refining search

Once users have developed their search concept and identified the keyword or keywords they need to search the database for records containing those terms. There are two ways users select to search the database: The first one is also the basic one that is to use the Find command to search for your keywords. The second method is refining search that is to select search terms from the Index or the Thesaurus and automatically have CDPIRS find them. The CDPIRS has Indexes an alphabetical list of all words in every field or several fields as you choose.

4.2. Narrowing a search

When users search retrieves too many records ,they can use any of the following techniques to narrow it:

operators (AND and NOT)
Field-Specific searching

4.3. Broadening search

When users search retrieves too few records ,there are several techniques they can use to increase the number of records retrieved, such as :

the OR operator
truncation
the Index
the Thesaurus

4.4. Display, print and download

Once users have retrieved records ,they can display them with the Show command ,they can mark certain records for later Printing or Downloading ,they can Print records, and Download records to save them on floppy or hard disk.

Users can output records in a variety of formats by changing the settings for the Show ,Print and

Download options.

5. Targets and Expectation

CDPIRS contains a large collection of standards, methods, policies, laws, programs, various of information and other data in this field. It contains information for the state and Chinese Disabled Persons' Federation in order to assist the government in mobilizing society for the development of disabled persons undertakings. Based on the information collected in CDPIRS, we provide service including :

Submitting reports of document statistic and data analysis to China Disabled Persons' Federation.

Providing report of document and data on special subject as needed.

Sending updated CDPIRS floppy and newsletter about CDPIRS to China Disabled Persons' Federation and its branches at provincial level.

Serving the organization and persons who are interested in the field of disabled persons.

Serving disabled persons directly.

Posters

Alternative pointing devices: some emphasises at designing of visual-independent HC-interface

Alexander Agranovskii, Tatiana Evreinova, Grigori Evreinov
SPECVUZAVTOMATIKA Design Bureau
Lab. for Designing of Information Image Systems
44/5, 13 Lenin St., 344038 Rostov-on-Don, Russia
E-mail: asni@ns.rnd.runnet.ru

Abstract

The paper are devoted to investigation of possible ways of designing of body-coupled manipulators. Described prototypes of alternative pointing devices were designed and tested. It is supposed that they are used jointly with audio feedback.

Keywords

feedback, panoramic adjustment, visual-independent interface, graphic manipulator

1 BACKGROUND AND PROTOTYPES

For estimation and correction of mouse operations an operator exploites not only afferent information of visual or unvisual receptors but also internal mental models as a coordinates system and comparison units: sizes of own body parts or of their relative displacement [Gurfinkel, V.S., Levik, Y.S. (1991)]. Using conventional pointing devices in absence / interruption of visual feedback an operator loses "a panoramic adjustment". It results to essential infringements of subjective impressions about environment objects. For blind users by the ideal tool could become input-output devices as much as possible using a metaphor of "mental notions of own body" (MNOB). Only in this case to an operator, irrespective of his physiological opportunities, the information about external objects and their relations will be adequately presented regarding individual coordinates. A becoming of intermodal MNOB is largely connected to a development of functional connections between tactile, kinesthetic and vestibular analyzers.

Capacitive joystick was performed with tactile markers (a clock-face metaphor of coordinates grid [Akamastu, M. et al (1995)]). Manipulator sizes corresponds to conventional mouse ones, and hand freely coveres a surface of joystick. The top movable part of the device is a hemisphere, radius of which is a little bit more than enclosed bottom hemisphere rigidly connected to the basis. A top hemisphere has perforations (holes) against of which on a bottom hemisphere the pins are placed. Functional buttons can be located on a bottom lateral surface of joystick. During positioning, the pins will be appearing in holes of appropriate sectors, a level of elevation and quantity of those is ample for recognition of position, type, state or of other attributes of external objects.

Figure 1 shows another type of manipulators designed on the same conception of MNOB. The input of graphic information is carried out via copying a trajectory of the pointer moving along a free surface of skin of the operator.

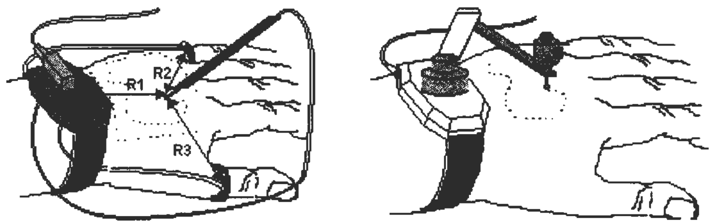


Figure 1 Graphic manipulators for pen-based computers using a skin feedback

Trajectories are calculated by triangle's method processing reactance's values between pointer, corpus and electrodes established on skin (on fingers). An electrical measurement of skin reactance is not deprived of common errors of contact methods. It complicates the processing of data. More higher resolution at low cost can be received using in manipulator construction potentiometric converters. The pointing accuracy with the help of tactile devices of such type surpasses a similar parameter of gestures interpreters based on force feedback (in pilot experiments it was not worse than ± 2 mm), as the range of signals of an accessible feedback is much wider (tactile and kinaesthetic cues are exploited simultaneously). Forehead, shoulder, hand, thigh or a surface of another body part can be used similarly, depending on an individual differential sensitivity and features of the device designing.

The described projects are supposed to be used jointly with audio feedback and speech synthesizer for blind computer users. But the last project can be effective as a communication device at more hard sensory pathology (deafblindness).

2 REFERENCES

Akamastu, M., MacKenzie, I. S., & Hasbrouq, T. (1995) A comparison of tactile, auditory, and visual feedback in a pointing task using a mouse-type device. *Ergonomics*, 38, 816-827.

Gurfinkel, V.S., Levik, Y.S. (1991) The conception of body model and a motion control. Intellectual Processes and Modeling. Moscow. "Science".

Strategy for Implementation of Information Technology in Health Care Sector for Developing Countries.

R. S. TYAGI [1]

Abstract

The application of Information Technology to Health Sector in the developing and 3rd world countries has not brought forward results with that pace as it has in economically developed countries. Experience provided by such an implementation at the All India Institute of Medical Sciences has been analysed and presented here to propose a strategy that is being adopted and implemented by the author for its hospitals and the other institutions in India. The past has been analysed and a proposal has been projected to provide sharing of resources by several hospotals in order to subsequently put the same to the benefit of all to achieve the goal – Health for all through IT applications.

SUMMARY

In the developing countries the application of Medical Informatics for the improvement of patient care, enhanced medical training, practices, and to provide exchange of information across the country has not been able to find its place so far. All India Institute of Medical Sciences (AIIMS) is a premier Medical Institution in the country having a large hospital of its own and four other associated centres having their own hospitals. AIIMS introduced the use of information technology in the information processing for the Patient Care System . It is a flexible system and allows for additions of new laboratories, new services and other facilities. It incorporates very special features for implementing the computerisation in a phased manner across the modules and across different locations for a module. This system was first started in 1989 meeting all its objectives. It was financed to its full in the initial period. Later it became difficult to meet the financial requirements for supporting the upgradation of the software , hardware and the application. The facts were analysed and after taking into consideration various factors it was proposed that solution is to be found by the industry and the user or a group of users by working together towards a common goal, though the interests of the two may be different. This requires responsibilities to be defined for each one of them and also shared between the two. The benefits accrued through such a joint venture shall go a long way. The All India Institute of Medical Sciences is one of such premier institutions in India that has put it to practice and the response from the industry have been extremely encouraging. The results of such a collaboration shall only be known after a passage of time. The information technology has a great potential and it holds the key to provide the proper and timely health care treatment .

[1] Department of Computer Facility, All India Institute of Medical Sciences, New Delhi, INDIA

The Poster Overview

Strategy for Implementation of Information Technology in Health Care Sector for Developing Countries.

Dr. R.S. TYAGI[2]

The poster is organised to project the subject matter under the following headings.

1. *Abstract*

2. Introduction to the subject matter

3. The computerised Patient Care System which was designed and implemented at the AIIMS.

4. The description of the functionalities of the following modules of the system.

4. 1. Central admission and charge collection.
4. 2. Wards
4. 3. Laboratories
4. 4. Operation Theatre
4. 5. Out Patient Services
4. 6. Emergency Services
4. 7. Blood Bank
4. 8. Billing
4. 9. Ancillary Services of Sterilisation and Crystalloids
4. 10. Diet for patients
4. 11. Roster for Doctors and Nurses, and
4. 12. Appointments

5. Analysis of facts leading to slackening and retardation of the system.

6. The identification of the problems responsible for hindrance in the implementation of the Information Technology in Health Sector.

7. The proposal to counter the problem and the responsibilities to be shared by Industry and the user.

8. The benefits through the joint venture of Industry and the user.

9. Conclusion

10. Diagramatic representation of certain existing systems.

---

[2] Department of Computer Facility, All India Medical Institute of Medical Sciences, New Delhi, INDIA.

Audio and Haptic Access to Math and Science - Audio graphs, Triangle, the MathPlus Toolbox, and the Tiger printer.

Steve Sahyun and John Gardner
Oregon State University, Corvallis, Oregon

and

Carolyn Gardner
Linn Benton Community College, Corvallis, Oregon

One of the greatest challenges for people with print disabilities has been the learning of, and access to, scientific information. Often, this information contains data, graphs, and equations which are represented in a highly visual manner. In an effort to make this data accessible to all individuals, the Science Access Project at Oregon State University has developed a number of new technologies that improve access to math and science for people with print disabilities. These projects include: audio representation of data graphs, the Triangle and MathPlus Toolbox programs, DotsPlus and the TIGER printer.

In the interest of creating an environment whereby students and professionals with print disabilities can have access to scientific information the Science Access Project at Oregon State University is working on a number of programs and research projects.

Audio representation of data, also known as sonification [1] of data, is a useful technique in providing quick, qualitative access to data sets.[2] Sonification has advantages over tactile printouts in that they require fewer resources, and are quickly and easily produced with appropriate software and standard computers. Recent studies [3] conducted by the authors explore the ability of students to interpret single-valued Y vs. X graphs created from tone plots. These graphs provide semi-quantitative non-speech auditory information on the value and slope of the function or data in question. It has been found that this technique shows promise as a method of displaying graphical information.

The TRIANGLE program demonstrates a self-voicing, self-Brailling computer application intended for access by blind people to math and science. This program acts as a mathematical scratch pad and permits reading, writing, and manipulating information using a linear notation convenient for both voice and Braille. The main features of Triangle include a text field, a table viewer, a graphing calculator, and an audio/tactile figure viewer. The text field allows for several independent editor text buffers with the ability to display and voice mathematical symbols. The table viewer allows easy access and display of elements in a table by moving from one cell to another. The graphing calculator can calculate mathematical functions and display data in a graphing window. The displayed graph allows for audio representation of the data for full access. Also, columns of data from a table can be plotted with the graphing calculator. The audio/tactile figure viewer provides voiced labeling of items in a previously constructed and annotated picture. The image is printed on a Braille printer and then placed on a touch sensitive device. The Triangle program is available in both Windows 95, and DOS versions.

The MathPlus Toolbox is a fully-accessible computer application intended primarily for teaching arithmetic and lower level math to children with learning or visual disabilities. It is a self-voicing program that provides work areas in counting, addition, subtraction, multiplication and fractions. It also includes a calculator and the ability for teachers to construct instructional lesson scripts that students can follow and work through.

DotsPlus is a device-independent typeface developed for blind readers that intends to encompass a large set of symbols in addition to standard braille. While computer technology has made it possible, often quite straightforward and inexpensive, to make words accessible to blind people,[4] making anything except words accessible remains a formidable challenge. Math equations, and figures such as maps, graphs, drawings, and charts that contain both graphics and characters (particularly those such as plus or equals that have no representation in standard literary braille) at unpredictable places pose a great challenge. [5,6] Some are almost impossible to make accessible to blind readers.

The DotsPlus tactile fonts [7-9] are designed to overcome these difficulties. In DotsPlus, literary braille symbols are used where possible, along with the raised representations of symbols not normally encompassed by braille, such as many mathematics and punctuation symbols, so the reader does not need to be familiar with math braille or computer braille codes. A one-page "cheat sheet" of the most common DotsPlus symbols contains enough information for a literary braille reader to be able to read almost anything in DotsPlus. Standard computer applications may be used, and little special training is needed by the person preparing a DotsPlus document.

The TactIle Graphics EmbosseR (TIGER) is a high resolution (20 d.p.i.) Braille printer capable of embossing text and graphics from standard Windows 95 computer applications. It offers unparalleled ease of use and functionality in that it can print files which contain both text and graphics directly from the applications from which one would normally print a document.

In addition to printing standard Braille fonts, the TIGER has been designed to print DotsPlus without difficulty. A document can be easily printed from Word, WordPerfect, or any other application that allows for font selection. The text is first formatted with either a six or eight dot version of the DotsPlus Courrier or Symbol fonts. These fonts contain the standard print characters, but the width of each character is set for proper layout when embossed as DotsPlus. The user then selects the print option from the word processor, chooses the Tiger printer from the list of available printers, and the TIGER embosses the DotsPlus page. Any vector graphic elements in the page are automatically embossed.

There is no difficulty in the combination of text and graphics other than layout considerations.

This research was supported in part by the National Science Foundation under grants HRD9452881 and HRD9353094.

1.Kramer, G. Ed.
(1994). Auditory Display, SFI Studies in the Sciences of Complexity, Proc. Vol. XVIII, Addison-Wesley.
2.Flowers, J. H., & Hauer, T. A.
(1995). Musical versus Visual Graphs: Cross-Modal Equivalence in Perception of Time Series Data. Human Factors, 37 (3), 553-569[2]
3.Sahyun, S. C. & Gardner J. A.
(1998) Testing the Equivalence of Visual and Auditory Graphs in Physics. Proceedings of the 1998 Winter American Association of Physics Teachers Conference, New Orleans, LA, January; http://www.physics.orst.edu/~sahyun/aaptw98/
4.Equal Access to Software and Information (EASI) http://www.rit.edu/~easi/
5.Gardner, J. A.
(1996) Tactile Graphics: An Overview and Resource Guide. Information Technology and Disabilities, 3 (4); http://www.rit.edu/~easi/itd/itdv03n4/article2.html
6.Schleppenbach, D.
(1996) Teaching Science to the Visually Impaired. Information Technology and Disabilities, 3 (4); http://www.rit.edu/~easi/itd/itdv03n4/article1.html
7.Gardner, J. A.
(1993) Dotsplus-Better than Braille? Proceedings of the 1993 International Conference on Technology and Persons with Disabilities, Los Angeles, CA, March; http://dots.physics.orst.edu/csun93.html
8.Barry, W. A., Gardner J. A. & Raman, T. V.
(1994) Accessibility to Scientific Information by the Blind: Dotsplus and Aster Could Make it Easy. Proceedings of the 1994 CSUN Conference on Technology and Persons with Disabilities, Los Angeles, CA, March; http://dots.physics.orst.edu/csun94.html
9.Preddy, M., Gardner J. A., Sahyun, S. C. & Skrivanek, D.
(1997) Dotsplus: How-to Make Tactile Figures and Tactile Formatted Math. Proceedings of the 1997 CSUN Conference on Technology and Persons with Disabilities, Los Angeles, CA, March.

VOICE
Giving a VOICE to the deaf, by developing awareness of voice to text recognition capabilities

Klaus Miesenberger, Gerold Wagner
Department of Computer Science, Johannes Kepler University, Linz, Austria

1. Background:

Deaf and hard of hearing people suffer from the lack of access to verbal communication. Help in reducing the gap between the deaf and the hearing world should be enforced. Automatic recognition of speech has to be considered as a powerful tool for closing this gap. Although voice to text recognition packages are already available and used for different purposes the community of the deaf and hard of hearing hesitates to use this new technology. This situation derives more or less from the lack of

- information about this technology
- knowledge about users needs
- studies of how and where to apply (feasibility studies)

2. Targets:

The VOICE Project intends to

- promote new technologies in the field of voice to text recognition (workshops, meetings, demonstration in practice)
- develop a demonstrator prototype
- set up test-applications (conversation, telephone, conferences, schools, TV)
- stimulate and increase the dissemination of information on the Internet
- act as an intermediary between the highly fragmented framwork of associations, companies, universities, schools, public administration,....
- specify and translate the users needs into technical specifications

Contact:

Modellversuch "Informatik für Blinde", University of Linz, Altenbergerstrasse 69, A-4040 Linz, Austria; Tel.: +43 732 2468 9232, Fax - 9322; E-mail: info@mvblind.uni-linz.ac.at

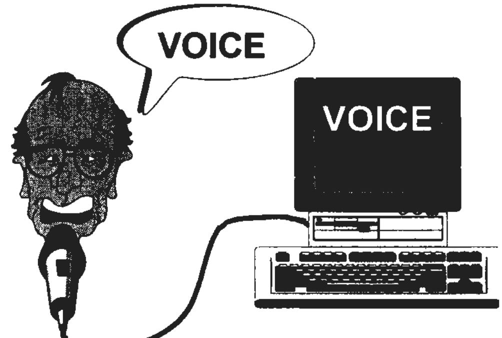
Acknowledgements

This work is being carried out in the context of the VOICE project (DE4104) TESTLAB project (LB 4003) funded by DG XIII of the European Commission under its Telematics Applications Programme.

Giving a Voice to the deaf, by developing awareness of voice to text recognition capabilities

Targets: promote new technologies of voice recognition
demonstrator prototype
test applications: conversation, telephone, conferences, schools, TV
dissemination of information
specify user needs

Information Technologies for People with Special Needs



consortium: JRC - ISIS - TP361, Ispra, Italy (main contractor)
Software Solutions, Milano, Italy
Associazione Lombarda Famiglie Audiolesi, Milano, Italy
Centro Comunicare è Vivere, Milano, Italy
Institut für Hör- und Sehbildung, Linz, Austria
Johannes Kepler Universität, Linz, Austria

Cooperations in Austria: ORF (Austrian Broadcasting Company), Vienna, Austria
OCG (Austrian Computer Society), Vienna, Austria
IBM Austria

picture **SENIOR CITIZENS** picture

SENIOR CITIZENS ON THE NET

a groupware for bridging generation gaps

SeniorOnline project

Organisation - who we are?

Picture and text

Objectives - what we want to achieve?

Picture and text

Research Methods - how we work?

picture and text

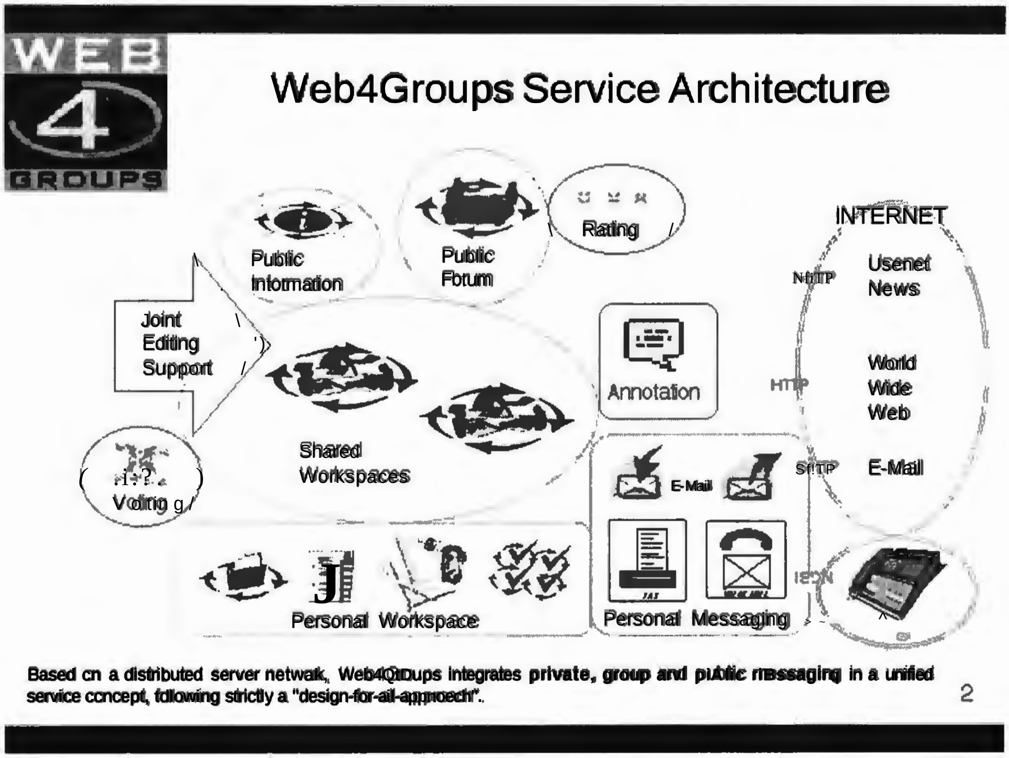
Web4Groups - groupware

How it came about?

Text

How we intend to develop it?

Text



Feedback booklet
for your comments

How to contact us

SeniorOnline is a project with the goal of stimulating elderly people to use information technology and the Internet. The project is funded by the European Union Telematics Applications Programme and there are partners from Italy, Sweden, Austria, and Ireland involved in it.

Background

The elderly are the fastest growing segment of the population in Europe, and it is estimated that their numbers will steadily increase towards the 21st century. At the moment, nearly a quarter of the population within EU (73,000,000 people) is 55 years or older. This development will create several new challenges for the society. How will the special needs of the elderly in the areas of leisure, health care, information, housing, etc. be met by society?

Many of the European governments are taking action in order to promote the use and knowledge of information and communication technology among the younger generations. This far the elderly have been forgotten. This is why SeniorOnline project is having the elderly in focus.

Objectives

SeniorOnline will provide simple access to the Internet for the elderly. The project will identify and assess the user needs and requirements and apply these to the technology earlier developed in the EUtelematics project Web4Groups. Basic to the project is:
1) The elderly is not a homogenous group and we need to identify the different needs of the different groups of elderly. However, the main principles of the software provided should be the same for all users in order to guarantee easy communication.
2) In order to prevent further fragmentation of the market, SeniorOnline will be based on existing technologies which will be further developed when necessary.
3) An assistive technology approach is specifically based on the elderly and their needs.

The description of the project

SeniorOnline consists of nine working packages: project management, external co-ordination and standards, active promotion among elderly people, user needs, functional specifications, demonstrator development, verification, demonstration, dissemination and exploitation. The project runs from 1998 to the beginning of 2000.

The Product

The Web4Groups is a group communication service. The innovative aspect of the Web4Groups service is that it will bring the full potential of a state-of-the art group communication tool to the Internet. It is designed to handle millions of users on thousands of interconnected servers. The focus of group support is on exchanging messages in forums or shared workspaces. Beside other advanced features, Web4Groups will support the collection and processing of votings and ratings.

Web4Groups allows any Internet user to join or start a public forum or a closed workspace, in order to exchange messages or multimedia documents (like word processing files or calculation tables). Objects are users, activities or messages. They are accessible not only through the WorldWideWeb, but also through e-mail (where they could replace list servers for mailing lists) as well as through fax or telephone.

Contact:
Sirkku Mānnikkö
Dept of Computer and Systems Sciences
Stockholm University/KTH Technical University
Electrum 230
S-164 40 Kista /SWEDEN
email:sirkku@dsv.su.se

Mechatronics Education for People with Physical Impairments (PPI)

Iwao Kobayashi[1], Toshiharu Nakazato[2], Sachiko Shiratori[3] and Masatoshi Iwamoto[4]

Abstract

We report our experience of teaching mechatronics for people with physical impairments (PPI). We explain the kit of the autonomic and a differential drive robot named "Bontenmaru" designed and developed as a teaching material. The result of questionnaire showed that most of PPI are interested in mechatronics and the class was successful.

1. Introduction

Mechatronics, the compound word of mechanics and electronics, is an engineering discipline for handling machines with circuits such as robots. We have taught classes on mechatronics for about 250 young students including people with physical impairments (PPI). We believe mechatronics is useful for PPI's independent living because in their everyday life they have to operate many products that are works of mechatronics, like electric wheelchairs. In addition, we expect that the development of PPI's software skills, that is gained through the mechatronics classes, is a way to increase their job opportunities in society. In this paper, we report on the educational materials we developed and on our teaching experience.

2. Educational materials

In 1996 we developed the kit of the autonomic and differential drive robot named "Bontenmaru" as a teaching material for a course on mechatronics aimed at young students. In order to design the autonomic robot for educational purposes, we designed the kit having the following five features:

1) Motor driver: Bontenmaru has a pair of rear wheels driven independently by two DC motors through a gear box.
2) Sensor: Bontenmaru can recognize obstacles in its forward direction by detecting the reflected IR light emitted by a pair of photo-diodes.
3) On board computer: Bontenmaru is controlled by a one-tip microcomputer PIC16C84 (Microtip Technology Inc.).

---

[1] Faculty of Software and Information Science, Iwate Prefectural University, Iwate, Japan, iwan@iwate-pu.ac.jp
[2] R & D Group, Scientia Co. Ltd., Sendai, Japan, nakazato@scientia.co.jp
[3] Miyagi Branch, Japan Muscular Dystrophy Association, Sendai, Japan, cyf03607@niftyserve.or.jp
[4] Faculty of Engineering, Tohoku Gakuin University, Miyagi, Japan, masa@cope.tohoku-gakuin.ac.jp

4) Programming language: The Japanese-based language "Makimono" was specified and developed for Bontenmaru and it allows Japanese children, who have poor literacy of English, to code down the control program on Microsoft Windows 95. The instruction set consists of moving commands such as "turn to the right" and condition branches such as "if the left detected then".
5) Remote control: As the IR detector is the same as those used in TV set, Bontenmaru can be remote-controlled by TV keybuttons. This helps Bontenmaru's operation by PPI who can not put it on the floor.

3. Management of the classes for PPI

Two classes for PPI, 2 hours each at one week interval, were held in National Nishitaga Hospital. The first class was on hardware assembling of Bontenmaru and the second was on coding the control program in Makimono. Most of PPI participants (denoted students) were patients with muscular dystrophy. 13 among 24 students were using computers and their age structure consisted of 14 in teens, 8 in twenties, one in forties and one unknown. Twenty-four volunteers (denoted teaching staff) had skills of computer use. After the class, questionnaire was conducted on both students and teaching staff to evaluate the class. Supplemental hearing investigation was carried out on students and nursing staff.

4. Results and discussion

96% of the students answered that they enjoyed classes, 100% understood the assembling Bontenmaru, and 95% got Makimono. Judging from the results, Bontenmaru excited and evoked their interest in the mechatronics. However, the students needed much help during the classes because 42% could not handle the screwdrivers nor nippers, and 21% could not finger keyboard of computer for program development. As 71% students could not put Bontenmaru on the floor and 58% could not connect/disconnect the Bontenmaru with ROM writer, most of PPI could not test their program by themselves. Then we prepared a special table for test trials and a pair of modular connectors which can be handled easily.

These results indicate that the classes have been successful, but that we do need to examine deeper a number of issues, like the design of educational materials for PPI and devising better ways to support the students during the classes such as improvements of a special table or a connector. Taking these experiences into account, Promotion Association of Mechatronics Education for Children is developing the next Bontenmaru.

We appreciate Promotion Association of Mechatronics Education for Children, National Nishitaga Hospital, Miyagi Branch, Japan Muscular Dystrophy Association and those who participated the classes for their support and cooperation. One of the authors (IK) thanks Mr. E. P. Duarte Jr. for his English correction.

REMOTE-CONTROLLED COMPUTER MOUSE FOR PHYSICALLY DISABLED CHILDREN

Terry Donnelly, James Windmill

Abstract
The use of a laser beam to move the screen cursor via a number of photo cells has been investigated, this proved difficult for some children to control. A later, and improved, system uses a pair of accelerometers to sense head movement which is then interpreted by digital signal processing (DSP) to move the cursor. Wrist-mounted tilt switches operate as mouse-select buttons. Future work will investigate the application of sensitive magneto-resistive (MR) sensors to detect limb movement. This will be a 'wireless' system but a significant problem will be the intensive, real-time computation required.

1. Introduction

Many children born with Cerebral Palsy have no power of speech and little control over their limbs. They require an effective means of interfacing with computers as soon as possible in their life so that they may participate fully in the educational process. Keyboard usage is usually out of the question and many cannot use a conventional mouse. The requirement, therefore, is for a cursor-controlling device capable of being operated by a part of the body, such as the head, over which the child has some control.

2. Modulated laser

This device was designed for a child who only had control over her head movement [1]. A small laser diode (taken from a laser pointer) is mounted on the front of a headband. It is powered from a small, battery-operated, 2kHz modulator to eliminate the effects of ambient light. The laser beam is directed at one of five photo-cells situated close to the computer screen. The photo-cells generate mouse-compatible clock signals when the laser beam impinges on the active part of the cell. Four of the cells represent *x,y* directions plus a fifth for select. The device proved useful but the child experienced difficulty keeping the beam focused on a photo-cell whilst watching the screen. A simpler system was required.

3. Accelerometer

The problems associated with the laser-beam mouse were taken on board in the development of a device which uses two accelerometers placed at right angles to each other. Rather than measure acceleration due to motion, the 'acceleration' due to Earth's gravity is measured. Used in this mode the accelerometers provide an output voltage proportional to tilt in the range ± 90 degrees. This output is applied to a DSP platform for software interpretation and generation of mouse-compatible signals. The select function is accomplished using tilt switches which can be attached to any limb. The accelerometers are small and light enough to be mounted on any part of the body which, when tilted in the *x,y* directions, controls the cursor.

One advantage is the freedom to tailor the system to the user via the software. For instance, to enable the user to hold the cursor stationary at the point of select a 'deadzone' was programmed in to prevent minor movement at zero degrees tilt. This can be varied to allow for the extent of involuntary movement of a potential user.

4. Magneto-Resistive (MR) Sensor: future work

The accelerometer (and laser beam) is an active device and requires a power supply and interface circuitry. Also, the user must be wired to carry the signals to the DSP platform. This inhibits extending applications to other areas, such as computer recognition of sign language, where a very small, passive motion/position detector is desirable.

It is planned, therefore, to investigate the application of Magneto-Resistive (MR) sensors [2]. These, solid-state, sensors measure magnetic fields and are very sensitive. Small, yet powerful, permanent magnets will constitute signal sources attached to parts of the body the position and motion of which is to be measured. The MR sensors, and associated processing electronics, can be located several meters from the user thus allowing freedom of movement.

A fundamental problem to be solved is the determination of the position and attitude of each magnet from the composite magnetic field detected by the MR sensors. The theory of superposition, which states that the composite field is the sum of the fields due to each magnet deployed separately, will be investigated. Other problems to be solved will be the elimination of the earth's magnetic field and the need for rapid, real-time computation of complex field quations.

5. Conclusions

The modulated laser device has been in use for three years. It has proved very useful but the user finds difficulty in concentrating on the position of the screen cursor whilst aiming the beam at the photo-cells: the accelerometer offers a solution to this problem. Also, in addition to interpreting co-ordinate movement and select functions, the accelerometer offers scope for interpreting key words or the alphabet thus opening up possibilities of access to the world-wide web.

The next phase of the work is to investigate the use of MR sensors. These will permit the use of small, passive, signal sources which will extend application, for instance, computer recognition of sign-language.

6. References

[1] DONNELLY, T., et al. Laser-Operated Mouse for a Physically Disabled Child, presented at IEE Colloquium: Computers in the Service of Mankind: Helping the Disabled, London, March 1997.

[2] MAPPS, D. J., Magnetoresistive Sensors, Sensors and Actuators, A Physical. A 59 (1997) 9-19, Elsevier. 1997.

Index of Contributers

INTERNATIONAL FEDERATION FOR INFORMATION PROCESSING (IFIP)

IFIP was founded in 1960 under the auspices of UNESCO. Its basic aims are to promote information in science and technology by fostering international cooperation, by stimulating research, development and the applications of information processing in science and human activity, by furthering the dissemination and exchange of information and by encouraging education in information processing. IFIP is a multinational federation of professional and technical organizations concerned with information processing. Currently it has 44 Full National Members, 3 Corresponding Members, 1 Associate Member and 11 Affiliated International Organizations representing all regions of the World.

IFIP's flagship event is the World Computer Congress, currently held biannually. In addition there are major international conferences and other events organized by IFIP's Technical Committees. Annually, IFIP is involved in the organization of over 65 events leading to some 30 to 40 IFIP books.

Preparations are under way for the 16th IFIP World Computer Congress, which will be held from 21 to 25 August 2000 in Beijing, China.

The core of IFIP's scientific work is carried-out by its 12 Technical Committees (with over 70 working groups):

TC 1 FOUNDATIONS OF COMPUTER SCIENCE

- WG 1.1 Continuous Algorithms and Complexity
- WG 1.2 Descriptional Complexity
- WG 1.3 Foundations of Systems Specifications
- WG 1.4 Computational Learning Theory
- WG 1.5 Cellular Automata and Machines

TC 2 SOFTWARE THEORY AND PRACTICE

- WG 2.1 Algorithmic Languages and Calculi
- WG 2.2 Formal Description of Programming Concepts
- WG 2.3 Programming Methodology
- WG 2.4 System Implementation Languages
- WG 2.5 Numerical Software
- WG 2.6 Database
- WG 2.7 User Interface Engineering (WG 13.4)
- WG 2.8 Functional Programming
- WG 2.9 Software Requirements Engineering

TC 3 EDUCATION

- WG 3.1 Informatics Education at the Secondary Education Level
- WG 3.2 Informatics Education at the University Level
- WG 3.3 Research on Education Applications of Information Technologies
- WG 3.4 IT Professional and Vocational Education in Infomiation Technology
- WG 3.5 Informatics in Elementary Education
- WG 3.6 Distance Learning
- WG 3.7 Information Technology in Education Management

TC 5 COMPUTER APPLICATIONS IN TECHNOLOGY

- WG 5.2 Computer-Aided Design
- WG 5.3 Computer-Aided Manufacturing
- WG 5.4 Industrial Software Quality and Certification
- WG 5.6 Maritime Industries
- WG 5.7 Integrated Production Management
- WG 5.10 Computer Graphics & Virtual Worlds
- WG 5.11 Computers and Environment
- WG 5.12 Architectures for Enterprise Integration

TC 6 COMMUNICATION SYSTEMS

- WG 6.1 Architecture and Protocols for Computer Networks
- WG 6.2 Broadband Communications
- WG 6.3 Performance of Communication Systems
- WG 6.4 High Performance Networks
- WG 6.6 Network Management for Communication Networks
- WG 6.7 Intelligent Networks
- WG 6.8 Wireless Communications
- WG 6.10 Photonic Communication Networks

TC 7 SYSTEM MODELLING AND OPTIMIZATION

- WG 7.1 Modelling and Simulation
- WG 7.2 Computational Techniques in Distributed Systems
- WG 7.3 Computer Systems Modelling
- WG 7.4 Discrete Optimization
- WG 7.5 Reliability and Optimization of Structural Systems
- WG 7.6 Optimization-Based Computer-Aided Modelling and Design
- WG 7.7 Stochastic Optimization

TC 8 INFORMATION SYSTEMS

- WG 8.1 Design and Evaluation of Information System
- WG 8.2 The Interaction of Information Systems and the Organization
- WG 8.3 Decision Support Systems
- WG 8.4 Office Information Systems
- WG 8.5 Information Systems in Public Administration
- WG 8.6 Transfer and Diffusion of Information Technology
- WG 8.7 Informatics in International Business Enterprises

TC 9 RELATIONSHIP BETWEEN COMPUTERS AND SOCIETY

- WG 9.1 Computers and Work
- WG 9.2 Social Accountability
- WG 9.3 Home Oriented Informatics and Telematics
- WG 9.4 Social Implications of Computers in Developing Countries
- WG 9.5 Applications and Social Implications of Virtual Worlds
- WG 9.6 Information Technology Misuse and the Law
- WG 9.7 History of Computing

TC10 COMPUTER SYSTEMS TECHNOLOGY

- WG 10.1 Computer-Aided Systems Theory
- WG 10.3 Concurrent Systems
- WG 10.4 Dependable Computing and Fault Tolerance
- WG 10.5 Design and Engineering of Electronic Systems
- WG 10.6 Neural Computer Systems
- WG 10.7 Microsystems

TC 11 SECURITY AND PROTECTION IN INFORMATION PROCESSING SYSTEMS

- WG 11.1 Information Security Management
- WG 11.2 Small System Security
- WG 11.3 Database Security
- WG 11.4 Network Security
- WG 11.5 Systems Integrity and Control
- WG 11.8 Information Security Education

TC12 ARTIFICIAL INTELLIGENCE

- WG 12.2 Machine Learning
- WG 12.5 Knowledge-Oriented Development of Applications

TC13 HUMAN-COMPUTER INTERACTION

- WG 13.1 Education in HCI and HCI Curricula
- WG 13.2 Methodology for User-Centered System Design
- WG 13.3 Human-Computer Interaction and People with Special Needs

For further detailed information on IFIP and its activities please visit the IFIP Web site at http://www.ifip.or.at

or contact the IFIP Secretariat:
Hofstrasse 3, A-2361 Laxenburg, Austria
Tel. +43 2236 73616,
Fax; +43 2236 73616 9,
e-mail: ifip@ifip.or.at

The following volumes of Proceedings of the six other Conferences of the XV. IFIP World Computer Congress are also available:

Global IT Security (SEC '98)
ISBN 3-85403-116-5

Fundamentals – Foundations of Computer Science
ISBN 3-85403-117-3

KnowRight '98 – 2nd Int. Conference on Intellectual Property Rights and Free Flow of Information
ISBN 3-85403-119-X

Teleteaching '98 – Distance Learning, Training and Education
ISBN 3-85403-120- 3

Telecooperation
ISBN 3-85403-121-1

IT & KNOWS – Information Technology and Knowledge Systems
ISBN 3-85403-122-X

All paper submissions are also disposable on a CD-ROM.

http://www.ocg.or.at/sr.html

f j

Kj

1
isT-

ISBN 3-85403-118-1 Österreichische Computer Gesellschaft