Integration of Adaptations for People with Special Needs

Hine N.A, Beattie W., Gordon I.A.S., Arnott J.L., McKinlay A.

MicroCentre, University of Dundee, Dundee, UK.

Spaepen A., Wauters F.

Katholieke Universiteit Leuven, Laboratory of Biomedical Kinanthropology, Tervuursevest 101, Leuven, Belgium

e-mail: nhine@mic.dundee.ac.uk

Abstract. The task of the IPSNI II project has been to investigate the accessibility of multimedia telecommunications services for people with disabilities. Where problems were found, the project was to develop and demonstrate examples of hardware, software and service adaptations solutions that show how accessibility could be improved. In doing so, the project has not only been able to utilise the flexibility and power of the intelligent broadband infrastructure, but has also raised a number of issues that are of interest to service providers and network managers. Some of these issues will be discussed in this paper in the context of two adaptations developed by the project. The principle emphasis of the paper is to illustrate some of the practical service provision and management questions that are raised when service adaptation is required for people with special needs.

1. Introduction

The use of a telecommunications service involves a number of distinct steps or tasks. Each task involves actions on the part of the user to perform a variety of actions. Examples of these actions include the manipulation of the terminal hardware user interface, negotiation of the terminal, network and service access procedures, and interaction with the information being exchanged.

This basic analysis simply lists a number of user procedures that are often performed without a second thought. For a user with disabilities, however, each task may become a barrier preventing access to the information being transported by the service. [1] Obvious examples include difficulty dialling using a keypad, or difficulties using a audio telephone service because of deafness.

This project considered three dimensions to the problem of access to services: The abilities or disabilities of the users, the set of services that they may wish to employ and the tasks that they may be seeking to accomplish. A series of mini-trials in the first year of the project allowed a small set of concrete problems to be identified, and adaptations have been developed to address these problems.

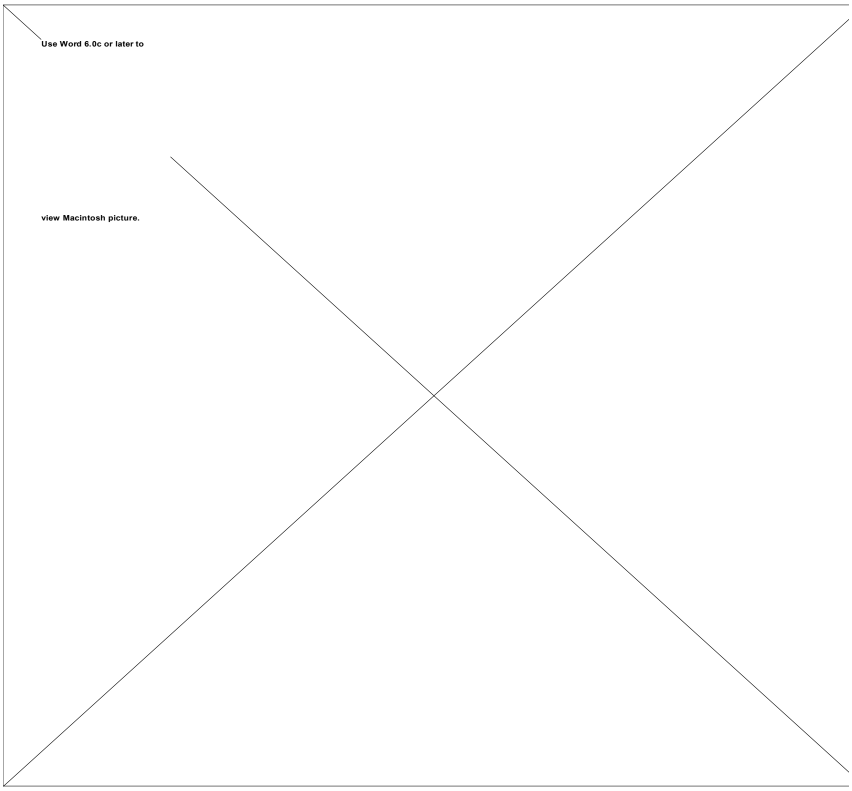

Fig. 1. Barriers to Information Exchange

The work of the project has been to develop the adaptations in the laboratories of the various partners, and then to integrate them into a Sun UNIX based parent platform at CSELT in Turin, or a PC based parent platform at VTT in Helsinki and Tampere. Both these platforms are connected by ATM networks. These platforms then serve as the host to demonstrate how a variety people with disabilities can gain access to a range of broadband multimedia services across an ATM network.

Two specific adaptations will be discussed in this paper as they allow the principle issues for service provision and network management to be highlighted and illustrated. The examples are the adaptation of a videophone service by the addition of a text telephone, further adapted with text prediction, and the addition of an eye gaze tracker to allow the user to drive the system and to manipulate information by eye gaze alone.

2. Adaptation Examples

2.1 Eye gaze tracking

People with severe motor disabilities, including those who have broken their necks or those who are in the advanced stages of a degenerative muscle or nerve disease, have major difficulties using conventional terminal input devices. Using the analogy introduced above, they are prevented from gaining access to the services at the first barrier.

One of the studies in the mini-trials in the first year of the projects was to investigate the use of an on-screen scanning keyboard. This is where the computer scans through each key on a graphical representation of the keyboard until the required one is reached. The user stops the scan by pressing a switch, either with their hand, or with a foot or even with their head. This device was tried because it is often the first choice of computer input device for users with severe motor impairments. The results of the mini-trials confirmed, however, that this type of input technique is very slow [2], and essentially impractical when text is to be exchanged in real time. For this reason, alternative devices that allow selections to be made directly from an on screen keyboard were considered, and an eye gaze tracker was selected by the partner at the Katholieke Universiteit Leuven.

Eye gaze tracking is a technique that has been used extensively in psychology research and in military systems. A variety of techniques have been employed to detect the movements of the eye and to determine the direction of the gaze. The one that is most frequently used in devices for disabled people is known as the corneal reflection method. In this technique, a low power infra red beam is directed towards the eye. The light beam is reflected off the retina, causing the pupil to glow, whilst at the same time, a bright spot is reflected on the cornea of the eye. The vector between the centre of the glowing pupil and the centre of the bright spot can be used to determine the direction of gaze.

The primary functional components of an eye gaze tracker are shown in fig. 2 below.

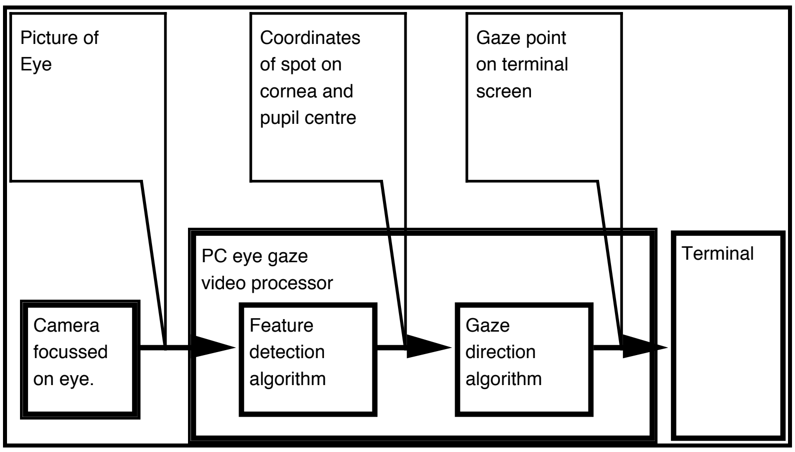

Fig. 2. The functional components of an eye gaze tracking system located locally

The system that is incorporated into the ATM test-bed system at CSELT uses a second camera to monitor the eye movements in addition to the one that is used for the videophone. The information from the camera is captured by a PC which processes 50 frames a second and passes the co-ordinates of the gaze point on the screen of the multimedia terminal to the terminal processor unit via a connection to the serial port.

This configuration is very expensive because it requires an additional PC to process the eye-gaze video data. In practice, many disabled people will have the least expensive terminal possible and will not have access to the computing power necessary to drive an eye gaze tracker. It is feasible that the required processing power could be located elsewhere in the telecommunications network. A major advantage of this approach is that people with severe impairments would be able to access services from conventional terminals rather than having to depend on specially adapted ones.

In addition, given that a multimedia terminal that can be used as a videophone already has a camera on it, it could be feasible to use the existing camera to monitor the eyes as well as provide the video for the videophone. This would require a high resolution camera if sufficient detail of the eye is to be visible within the whole view of the head. In this case, the functionality could be distributed as shown in figure 3.

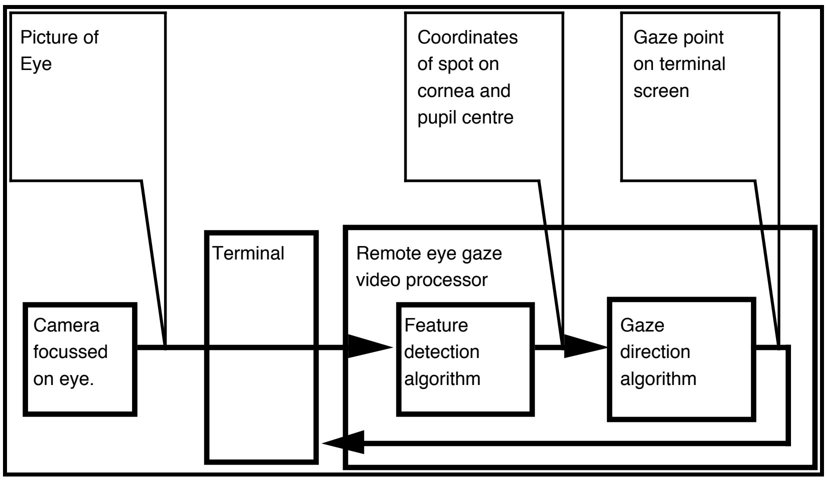

Fig. 3. The functional components of an eye gaze tracking system located remotely

This configuration would appear to offer an ideal solution to the need to allow users with severe impairments access to services via a low cost terminal. It is important to check, however, that the configuration is usable. A primary consideration might be delay.

A typical eye gaze tracker samples the eye camera every 20ms or 50 times a second. If 5 consecutive frames show the same gaze direction, a fixation is recorded, and the cursor is moved to the point on the screen being gazed at. This takes 100 ms. In order to select an object on the screen, gaze is maintained on the same point on the screen for between 500ms and 2s, depending on the experience of the user.

Whilst tracking and gazing in a general direction are tasks that the eyes are well suited to, directed gazing for prolonged periods is not. The eye muscles fatigue and small tremors in the eye muscles cause the system to ignore attempts to fix on a point.

For this reason, any delay in the system needs to be kept to a minimum in order not to prolong the time needed to register a selection.

In addition, if the network is imposing a delay on the processing of the data, there will be a lag in the movement of the cursor as it follows the movements of the eye. The user may be able to cope by tracking across the screen sufficiently slowly for the system to catch up, but this will require fine eye control, which again can be tiring. Alternatively, the user will jump back to the last cursor position and try and "pick-up" the cursor. This will send the user and the system into oscillation as they both try to keep up with each other. One soultion is to attempt to predict the path that the user is intending to track as they move their eye, and correct the path based on actual movements made. Further research is need to tackle these issues.

In a pure ATM environment with efficient packetting algorithms and small queuing delays at the switches, the response of the network should be sufficiently fast to be effectively real time. [3] In the evolutionary phase between the current technologies and widespread availability of ATM infrastructures, delays may be introduced at the boundaries between networks of different types, particularly across LAN, MAN and WAN boundaries and exchanges. These delays may be sufficient to make this configuration unworkable unless the remote processing power is located at a site that is geographically fairly close to the user's terminal with few network switches between them.

In addition to this issue, careful consideration will need to be given to the user procedures involved in gaining access to the eye gaze tracking support service, given that some of the essential functionality is located remotely and will need to be invoked before it is available to be used.

Telecommunications costs are another important issue. Even a low bandwidth service such as retrieval of information from a text data base is always going to involve an additional high bandwidth requirement as the video images are passed for remote processing. The bandwidth required for the video is in the order of 60 Mbits/sec. The current implimentation looks initially at the area of the eye image where the omportant features were last found. By using such a technique, most of the bandwidth requirement can be reducded to around 0.6 MBits/sec if the remote processor could call for seleced parts of the image. Lossy compression could not be employed as this would make the feature detection too innacurate.

2.2 Videophone, text phone & text prediction

The eye gaze tracking is an example of a solution to the first barrier encountered when trying to access a service. This section will concentrate on an example of a solution to the third barrier, that is, providing a communication channel when a medium is inaccessible. The example described here is being investigated by the partner at the University of Dundee. They are concerned with the situation where the user is unable to use a videophone because they have difficulty speaking. A possible solution is to add a text telephone service component to the videophone service. This adds little to the network load or management of the service so would appear to be an ideal solution.

The complication arises from the fact that many non-speaking people have an additional motor impairment that has the consequence that they are slow at typing. To counter this, the group at Dundee have added text prediction to the text telephone service. This assists the user by presenting a selection of words when the user starts to type a new word, based on the letters already typed and the words that the user tends to type using those letters. This can save the user up to 50% of the key strokes that would be involved in typing the same words without prediction, and can reduce the time involved in typing by about 20% [4].

The issues that become important from the point of view of the user are the availability of the predictive assistance, and the learning that the predictor achieves. The implementation of prediction in response to these issues may have implications for user procedures and the service provision.

The function of the predictor is to monitor the keys pressed by the user, and to suggest a list of words based on the key presses. If a new word is typed, it can be added to the lexicon being used by the predictor. When a word is used that has been suggested by the predictor, it rises in prominence in the suggestion list. The priority given to a word is based both on the recency with which it has been used and the frequency with which it has been used. [5]

When prediction is used in the context of telecommunications interaction, there is a possibility of it learning from the other parties in the conversation in addition to the user who is actually employing it. In a conference situation, it is important that the user has control over the parties that are contributing to the learning process. For example, American English has a number of different spellings from European English. The user will probably want to look at the contributions from each participant and then configure the predictor to learn from some participants and not others. Depending on where the text-telephone functionality and the prediction functionality are located, this may have implications for service functionality and service provision.

The two configurations that have been implemented in the project are shown in figure 4 and 5 below.

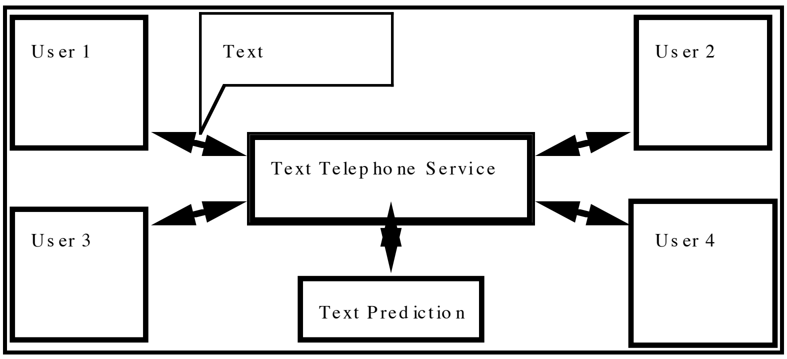

Fig. 4. Central Text Telephone Service with Text Prediction

In the configuration in figure 4, the prediction function is offered as part of the text phone service. Each person taking part in a text conference is offered the predictive assistance, and each user of the prediction has their own lexicon, which can learn from any or all of the participating users. Because it is tied to the text phone service, it is not available to any other service (e.g. text searching of a data base) or any other application running in the terminal.

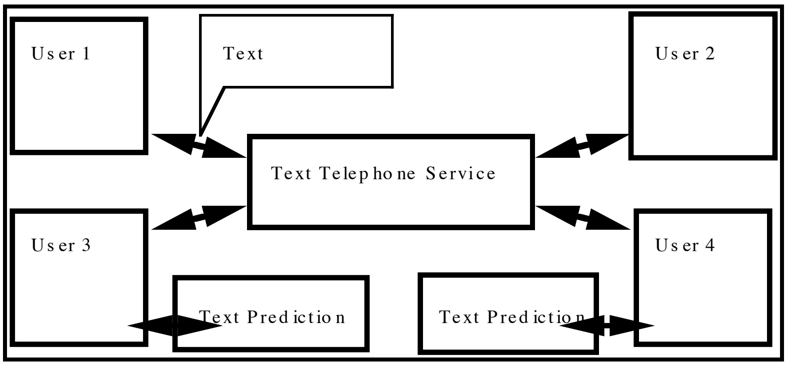

Fig. 5. Central Text Telephone Service with Terminal based Text Prediction

In the configuration shown in figure 5, the prediction functionality is provided as an application running on the terminal. In this case it is available to be used with any application running in the terminal or any service being accessed by the terminal. Because it is learning by monitoring the “keyboard buffer function” of the terminal, it cannot learn from text coming in from the network, but only form text entered by the user of the terminal.

An ideal solution would be a predictor that was able to learn from all text passing through the terminal. In this case, the user should be able to determine which sources are contributing input to the predictor. For this, the predictor should be able to identify the individual contributors in a text conference. This might be difficult if the text telephone is served from a central location as in figures 4 and 5, but possible if the text phone is an application running in each terminal that sends and receives text directly to and from each terminal in the conference.

This example has provided a platform to illustrate that the concrete implementation of a relatively simple service such as text telephone can have implications for the features that can be provided in an adaptation to assist a user with impairments.

3. Implications

The examples highlight some of the issues encountered by the IPSNI II project when it has sought to integrate an adaptation for people with disabilities. The project has had considerable freedom to explore the issues in a laboratory environment where emulated services are running on an open multimedia platform. A number of open questions remain, however, which become pertinent when consideration is given to adaptation that will be required on terminals built to allow access to real services running on the future broadband infrastructure. These are summarised below.

Terminal design:

Terminal equipment for the domestic and business environments will rapidly evolve into closed boxes designed and engineered to suit the target setting. In practice, this invariably means a piece of equipment that is difficult to adapt. The availability of intelligence elsewhere in the networks is an ideal approach to providing powerful adpations, provided that the basic user interface and ergonomics can be adapted where required. This will depend on standard interfaces being employed between hardware components and software components of the terminal. These standards should be adopted de facto or formally for even the cheapest equipment, as many people will not be able to afford the more expensive, better quality terminals.

Distributed intelligence

There are a number of parameters that could affect the usability of functions that take advantage of employing intelligence located remotely from the user's terminal. These include bandwidth availability, delay and synchronisation. In principle, these aspects are foreseen to be transparent to the user, but it is likely that in the crucial evolutionary phase to full ATM implementations, they will be important constraints. Because the adaptation of service access may change the behaviour of a conventional service quite considerably, the initial network and service management algorithms may have difficulty allowing the flexibility that could be required.

Alternative Services

It is clear from the work of the IPSNI II project that there will be some situations where it will not be reasonable to adapt a terminal or a service. In these cases the intelligent network provides unparalleled opportunities for the development of special services. An example is the provision of multimedia computer based interviewing services as a preliminary step in a discussion. This technique could find application in situations such as initial discussions prior to booking an appointment with a doctor or when arranging entitlement to social benefits.

4. Conclusions

The way people with disabilities use or gain access to services may be very different from an equivalent utilisation by able bodied people. A call involving a user with disabilities may involve extra media to compensate for the innaccesible information type, or it may have diffrent usage characteristic if a user is slow or has difficulties understanding or being understood. Service use by people with disabilities is therefore a useful way of validating the functionality of the infrastructure and management techniques for all users. The broadband telecommunications infrastructure holds tremendous promise for facilitating involvement in many areas of life for people with disabilities. For this promise to be fulfilled, it should have all the functionality necessary for any requirement to be met. The critical element necessary to ensure access for all users is that the functionality has been implemented sufficiently flexible for it to be utilised in novel ways when required. The IPSNI II project has sought to expose some of the underlying issues and implications so that they can be taken into account at the design phase of the infrastructure and service management procedures.

5. References

[1] Byerley, P. and Connell, S., “Integrated Broadband Communications: Views from Race, Usage Aspects”, Elsevier Science Publishers B. V. (North-Holland), 1992, pp. 201-230.

[2] “Consideration of Scanning Keyboard and Text Prediction in the context of access to Telecommunications Services” N. A. Hine, W. Beattie, A. McKinlay and J. L. Arnott, ISAAC: 6th Biennial Conference of the International Society for Augmentative and Alternative Communication, Maastricht, The Netherlands, 9-13 October 1994.

[3] M. de Prycker, “Asynchronous Transfer Mode: Solution for broadband ISDN”, Ellis Horwood Limited, 1991, pp. 63-69

[4] IPSNI II, “Access solutions for people with special needs to telecommunication”. Deliverable to the CEC from RACE Project R2009 IPSNI II, Workpackage (WP 1.4), CEC Deliverable Number R2009/IRV/WP1/DS/P/008/b1, 1992.

[5] Swiffin, A.L., Arnott, J.L., Pickering, J.A. and Newell, A.F., “Adaptive and predictive Techniques in a Communication Prosthesis”. Augmentitive and Alternative Communication (A.A.C.), Vol.3, No.4, December 1987, pp. 181-191.