

ARCHITECTURAL ISSUES GOVERNING THE IMPLEMENTATION OF A
MULTIMEDIA PORTABLE COMMUNICATION DEVICE

Nick Hine
Department of Applied Computing
MicroCentre, University of Dundee,
Park Wynd, Dundee,
Scotland
DD1 4HN
Tel: +44 (0)1382 344711
Fax: +44 (0)1382 345509
e-mail: nhine@mic.dundee.ac.uk

Abstract

The dream of non-speaking people is to have a powerful yet portable communication device that will enable them to communicate while they are mobile. This would allow them to communicate with other speaking or non-speaking people.

It is proposed to provide a user with a computer system as a platform for an assistive communicator that will help them in face to face conversations or when communicating in a group with other non-speaking people. In addition, they could use the computer to communicate with a remote user via a videophone or similar service. In all cases, the assistive system will help the user to:

- 1 Produce text (which might subsequently be spoken by the system),
- 2 Find, display and annotate pictures
- 3 Find and play video or sound clips

This paper describes such a novel assistive communication system for non-speaking people. Text chatting, picture display and annotation, and the presentation of video and audio clips are included in the architecture of the system. This will be supported by a powerful conversation tracking engine which will monitor the information exchanged and used in the conversations, and will predictively retrieve material from the data store that can be used later in the conversation.

Introduction

Telecommunications is becoming increasingly important in all aspects of life today. In business, in addition to conventional telephony, many people regularly use fax services. Studio and desktop based videophone services are now being used as an alternative to face-to-face meetings.

At home, the conventional telephone is being used to provide a wide range of services such as 24 hour banking and 24 hour access to insurance. More powerful multimedia services will become more prevalent as the number of cable TV installations increases. Other trends include a greater penetration of the Internet, increased use of personal computers for accessing information and for remote communications, and increased bandwidth available in telecommunications networks. All these trends point to a world where the use of telecommunications services will become essential in many aspects of daily life.

With this in mind, it is obvious that a user with a communication impairment will be at a severe disadvantage. They will encounter difficulties in participating in activities that most members of the population will take for granted and depend on.

The work of the CEC IPSNI funded IPSNI project (1990-1992) and IPSNI II project (1992-1995) demonstrated that it is technically feasible for a telecommunication terminal to be adapted so that non-speaking people can participate in conferencing/conversation services.

The primary information media in interpersonal interactions and therefore the associated conversation services is speech. As a substitute, text is often used in low bandwidth situations (for example text telephone facilities on Internet videophones). A logical step would be to propose text as an alternative media for people with speaking difficulties. This is not a realistic solution in many cases as a high percentage of non-speaking people have an additional motor impairment that makes typing difficult, or they have reduced language abilities. In Augmentative and Alternative Communication (AAC) devices symbol systems are often provided as an alternative to text. However, many non-speaking people, particularly adults, do not use these as they are not understood by many able bodied people. The more powerful symbol systems are effectively a language, and require the same degree of tuition and practice as traditional orthography.

Results from the IPSNI projects suggest that additional media such as pictures, audio or video may be a useful addition to an assistive communication system in order to highlight specific facts, or even to set a general context. If a communication device allowed communication parties to annotate pictures, attention can be drawn to detail within

the picture. This would allow quite specific information to be exchanged without needing to employ specific textual explanation or other language like constructs.

Research effort at the University of Dundee is currently concentrating on providing users with a portable computer based communicator with information presented using media such as pictures, video clips and audio clips as well as text. A selection of media items can be presented to the user, who can then select an item and present it to the other participants in the conversation. Additional emphasis can be provided by allowing the user to annotate pictures to highlight aspects of interest and relevance to the conversation. As the conversation progresses, a conversation tracking engine can analyse the direction that conversation is taking and using a prediction algorithm, present a new set of media items to the user.

This type of assistive service will require a powerful conversation analysis engine and the ability to recall information in a variety of media from a rich, and therefore extensive store. It should also be available on a low cost portable system that can be available whenever a disabled person needs to communicate. The service should also be available to operate on a telecommunications terminal alongside other telecommunications services such as a videophone. This is a demanding requirement, and it is unlikely that the facilities can be realised on a self-contained portable device. For this reason, an architecture is proposed for the services that depends on a client/server architecture, with a central server supporting remote clients. In order to ensure that a rich conversation can be supported, the conversation tracking engine, and the information stores will be located on central servers. Information will be uploaded from the servers using mobile phone or Wireless LAN links. The architecture will scale the behaviour of the service to accommodate the information transfer capacity of the data communication link.

This approach is related to the InfoPad [Narayanaswamy et al 1996] and PARCTAB [Want et al 1995] work. These investigate the provision of complex services on portable handheld devices, and demonstrate these as the user interface to services running on remote servers.

System Architecture

The conversation system consists of four key components. These are the Presentation Functions (in the mobile terminal), the AAC Presentation Function (in the mobile terminal), the Conversation Tracking Engine (in a remote server) and the Media Server (in a remote server).

The heart of the presentation functions is a text display area, or if used in conference with other terminals, a text telephone. Users can type material directly into the text phone. At the same time, however, a set of items of information in a variety of media are available to the user via the AAC presentation function. These can be selected from the AAC presentation function, and they are passed to the appropriate Presentation Function for presentation to the other conversation participants. This all takes place in a portable terminal.

The media items used are reported to the Conversation Tracking Engine. This then checks to see if this is a new user generated text or annotation or if it has been selected from the AAC Presentation Function. If it is a new media item, this item is passed from the generating Presentation Function to the other Presentation Functions taking part in the conversation. This takes place between the portable terminal and a remote server over the mobile link. Any additional portable terminals are also connected to the remote server via mobile links

If, however, it is an item selected from the AAC Presentation Function, the Conversation Tracking Engine will tell the other Presentation Functions to retrieve this media item from their local AAC Presentation Functions. Then the Conversation Tracking Engine determines the most appropriate media items that are most likely to be needed next in the conversation. It passes the identity of these items to the Media Store for retrieval and passing to the AAC presentation function.

All media exchanged during the conversation will be logged by the Conversation Tracking Engine. Following the conversation, the log will be analysed. Each piece of media will be checked, and its relationship with the other media used will be checked. This will be reflected in an updated entry in a dynamic list. In this way, items of media that are often used in the same conversation will be most likely to be predicted and available for use in subsequent conversations.

Two studies have been conducted to consider the validity of this approach. In the first study, a number of non-speaking users were provided with interfaces that allowed them to exchange information in a variety of media, including picture display and annotation. This study confirmed that for some users, picture annotation improved the rate at which information could be exchanged, and for all users it improved the "richness" of the information exchanged. It is clear that the implementation of the user interface of any system will need to reflect the fact that many non-speaking people also have physical impairments that can reduce their ability to handle a communication device. [Hine et al 1997].

The second study considered some of the technical aspects of the link between the server and the portable communicators. The results demonstrated that a link provided by a mobile cellular phone would be sufficient to provide a text based remotely running prediction service. The also demonstrated that a 3 Mbps wireless LAN link was functionally equivalent to an Ethernet fixed link. This type of link would be sufficient to provide a properly constructed multimedia communication service. [Beattie et al 1997][Hine et al 1998]

Conclusion

This work seeks to provide non-speaking people with a communication device that is both powerful and portable. It combines a number of advanced technologies, and takes into account the specific characteristics both of the users and of the technologies employed.

References

W.Beattie, N.A.Hine and J.L.Arnett., "Distributed Assistive Communication Devices For Non-Speaking Users." Proceedings of 16th International Symposium on Humans Factors in Telecommunications May 1997, pp 185 - 190.

Hine, N.A., Wilkinson, D, Gordon, I.A.S & Arnett, J.L. (1995b). An Adaptable User Interface To A Multimedia Telecommunications Conversation Service For People With Disabilities. In Proc. Interact '95, 1995, Chapman & Hall, pp394-397.

N.A.Hine, W.Beattie and J.L.Arnett., "Study of Picture annotation as a Means of assisting Non-Speaking People to Use Telecommunications Services." Proceedings of 16th

International Symposium on Humans Factors in
Telecommunications May 1997, pp191 - 198.

IPSNI ii (1995). CD-ROM Integrated Telecommunications For
People With Special Needs.

Narayanaswamy S., Seshan S., Amir E., Brewer E., Brodersen
R.W., Burghard F., Burstein., Chang Y-C., Fox A., Gilbert
J.M., Han R., Katz R.H., Long A.C., Messerschmitt D.G.,
Rabaey J.M., "Application and Network Support for InfoPad",
IEEE Personal Communications, vol. 3, no. 2, April 1996,
pp. 4-17

Want R, Schilit B.N., Adams N.I., Gold R., Petersen K.,
Goldberg D., Ellis J.R., Weiser M., "An Overview of the
PARCTAB Ubiquitous Computing Experiment", IEEE Personal
Communications, vol. 2, no. 6, Dec 1995, pp. 28-43