5-2 Barrier-Free on the Mobility and the Information for Visually Impaired People and Hearing Impaired People

OYAMA Shin'ya, IGI Seiji, NISHIMURA Takuichi, and YAIRI, Ikuko Eguchi

We have been developing Robotic Communication Terminals (RCT) which support the self-mobility of the elderly and disabled people. One of the terminals we developed is "user-carried mobile terminal" which gives the information such as the navigation to visually and hearing impaired people who can walk by themselves. In this paper, we introduce the animation system to show the sign language for hearing impaired people and the voice guidance system for visually impaired people with the infrared communication and AM radio communication.

Keywords
Self-mobility, Visually impaired people, Hearing impaired people, Animation of sign language, Infrared communication

1 Introduction

This research is focused on the issue of mobility, one of the most basic and indispensable of human activities. The aim of the research is to facilitate mobility for the elderly and disabled facing difficulty in moving about independently. Older people and those with impaired vision, hearing, and disabilities in the lower extremities face problems in the environmental recognition, physical actuation, and information access required for full mobility, and thus are challenged by a range of difficulties in self-mobility. To address these problems, we launched a research project to develop the "Robotic Communication Terminal (RCT)"[1], with the aim of facilitating self-mobility of the elderly and disabled.

As shown in Fig. 1, the RCT system comprises three types of terminals: an environment-embedded terminal, which monitors the mobility environment; a user-carrying mobile terminal, which carries a user that has difficulty in walking while monitoring the surrounding circumstances; and a user-carried mobile terminal, which provides information including local navigation data for users who can walk by themselves. The terminals each have different roles, and by communicating with each other, they provide coordinated assistance to the user in recognition, physical actuation, and information access, and generally facilitate user mobility in urban and other areas.

Among these terminals, the authors have been working on the development of the user-

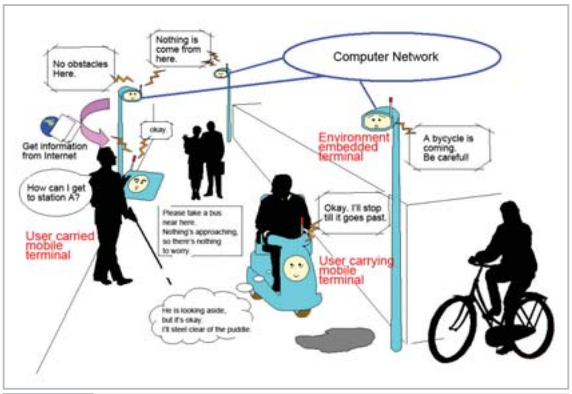

Fig. 1 *Schematic diagram of support for self-mobility through the RCT system*

carried mobile terminal. We have conducted research and development of a method of presenting information to the hearing impaired based on computer graphics (CG) sign language animation displayed on a mobile information terminal and an area guidance system to promote self-mobility among the visually impaired. This paper provides brief outlines of these systems and actual application experiments performed with the cooperation of disabled volunteers.

2 Information presentation by sign language animation in mobile environments

Hearing impairment is sometimes referred to as an "information impairment", attesting to the significant difficulties faced by the hearing impaired in daily communications, particularly when traveling from one place to another. To help facilitate communications between the hearing impaired who use sign language in daily life (for simplicity, referred to as the "deaf" below) and those that are unimpaired, we have developed two techniques. One technique consists of a system of sign-language recognition[2] through image processing and conversion of the recognized data into voice information. The other technique[3] involves recognition of voice data and subsequent generation of sign language using computer graphics (CG) tools. The sign-language recognition technique itself presents many problems that remain to be solved before practical application. However, the technique of generating sign language is already in a stage approaching practical application.

The recent expansion of wideband information and communication networks as well as the spread of mobile information terminals have made it possible to provide services involving various types of animation. Here, we describe a technique of generating sign language and a technique for the presentation of CG sign-language animation displayed on a mobile information terminal (PDA) with the aim of assisting deaf people in movement.

2.1 Sign language animation generation system

There are two types of techniques for generating sign-language animation: those that are based on motion capturing — in other words, techniques that import the movements of sign language by deaf people into the computer and reproduce these movement as animation[3]— and techniques based on motion primitives, which encode and synthesize the movements of sign language strictly through programming[4]. Techniques based on motion capturing can generate realistic sign language but each requires massive data acquisition and correction processes. On the other hand, techniques based on motion primitives offer the advantage of convenience in that these methods generate sign language simply through a computer program; however, the generated animation lacks the realism of motion-capturing techniques.

Figure 2 shows the configuration of the developed system. It is reported that there are approximately 4,000 sign language words. When a Japanese sentence is input, the system performs a morphological analysis, extracts the words required for expressing the sentence in sign language, and retrieves the corresponding sign language motion data from the built-in word database. Then, the system interpolates movements between the words and generates a sign language sentence. The system stores 1,500 words previously prepared by motion capturing. We selected words from sentences relating to situations in which deaf

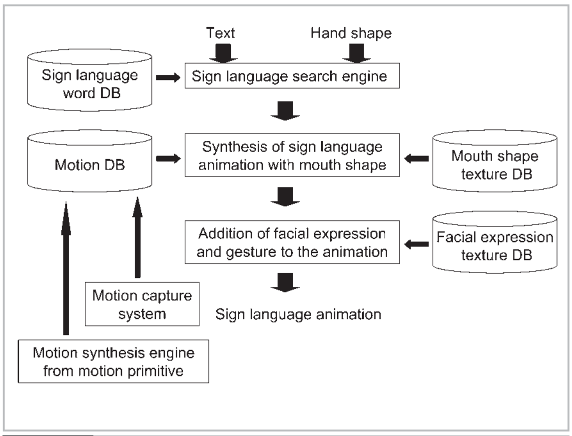

Fig.2 Configuration of sign-language animation generation system

people are often inconvenienced, such as in hospitals, banks, and traffic. As words not stored in the database may be required when generating sign language, the system is constructed to support the importation of CG data based on motion primitives. The system can synthesize not only the motion of the hands and fingers but also facial expressions, gestures, and mouth forms, which are also important elements of sign language.

2.2 Local information delivery to moving users

2.2.1 Local information delivery system

Figure 3 shows the mechanism behind the automatic delivery of information to deaf people moving while using PDAs. As an example of practical use, we assumed a scene in which the system is delivering information relating to a museum visit. The system consists of (1) an information delivery server, (2) a positional information detection server, and (3) a wireless LAN system. Each PDA is equipped with an infrared tag that transmits unique ID information. The positional information detection server continuously determines which PDA (and owner) is in front of which exhibit via the infrared sensor. The PDA is installed with middleware that receives positional information from the positional information detection server and enables the information delivery server to deliver the corresponding information. As the owner of the PDA moves, the information on the corresponding exhibit is displayed on the PDA in sign-language animation. This animation uses 208 × 180-dot frames at a rate of 20 frames per second, and is delivered via 80-kbps streaming.

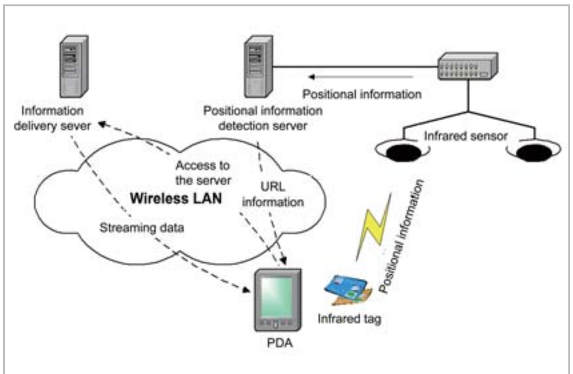

Fig.3 *Method of positional detection and information delivery according to position*

2.2.2 Evaluation experiment and discussion

We performed an experiment to assess recognition of the sign-language animation displayed on a PDA. The examinees consisted of seven people who were either congenitally deaf or who had lost their hearing by the age of three. Table 1 shows details on each examinee. In the evaluation we provided each examinee with a PDA and asked each to operate the PDA on his or her own three times, in order to view the animation of signs and sentences and then asked that each write an evaluation.

(1) Word evaluation experiment

Sixty words based on motion capturing and 30 words based on motion primitives were evaluated, for a total of 90 words for both types together. Table 2 shows the results. The accuracy rate was 91.6 percent for words based

Table 1 *Examinee information*

Examinee	Age	PC experience	E-mail experience	PDA experience
A	28	Yes	Yes	No
B	20	Yes	Yes	A little
C	25	A little	No	No
D	57	Yes	No	No
E	39	Yes	Yes	No
F	26	Yes	Yes	No
G	29	Yes	Yes	No

Table 2 *Evaluation results of sign language word and sentence animation*

Examinee	Accuracy rate of words based on motion capture (%)	Accuracy rate of words based on motion primitives (%)	Accuracy rate of sentences (%)
A	91.7	80.0	81.8
B	95.0	63.3	100
C	98.0	63.3	54.6
D	75.0	40.0	54.6
E	90.0	50.0	45.5
F	91.7	70.0	63.6
G	100	80.0	100
Average (%)	91.6	63.8	71.4

on motion capturing and 63.8 percent for those based on motion primitives.

(2) Sentence evaluation & recognition experiment

Eleven sentences extracted from descriptions for museum exhibits were evaluated in the experiment. The longest sentence contained 17 words, while the shortest one contained six. The average number of words contained in a sentence was 12. Table 2 shows the results. The sentences were evaluated on a four-grade scale: “correct” (the meaning of the sentence was completely understood), “almost correct” (one to three words were incorrect but most of the meaning of the sentence was understood), “half understood” (about half of the sentence was understood), and “erroneous” (including blanks). On average, 44.2 percent of the sentences were completely understood, and 27.2 percent were almost completely understood, for a total of 71.4 percent completely or almost completely understood.

(3) Discussion

The accuracy rate was slightly lower for words based on motion primitives. Comparing the results with those of an evaluation experiment for words and sentences used in daily life displayed on an ordinary PC screen, in which the corresponding accuracy rates were 93 percent and 92.5 percent, respectively[3], the accuracy rate values were lower in this experiment. This is probably due to the academic nature of the content used in the experiment — descriptions of the art exhibits — and also due to the limited size of the PDA screen. In a past experiment to evaluate primitives[5], we obtained a result that adding mouth shapes improved accuracy. However, in this experiment, the examinees could barely identify the mouth shape on the small PDA screen and thus this feature did not help improve accuracy.

Viewing comments by age, we observed that older people noted that the screen was small; responding to the needs of the elderly in this area thus remains a future challenge. On the other hand, younger people commented that the device was easy to use. Despite some age-related differences, our results generally show that streaming images of CG animation sign language to a screen the size of a PDA can in principle be recognized, providing a viable means of communication to deaf people in motion, in accordance with their positions.

3 Mobility support for the visually impaired

3.1 Mobility support for the visually impaired based on CoBIT

Although many systems have been proposed to support self-mobility of the visually impaired, most of these systems rely on highly functional terminals such as PCs and PDAs[6][7]; one may conclude that visually impaired users will face significant difficulties in operating and maintaining such devices.

On the other hand, a system based on terminals that are easy to handle and that can provide the information required for movement can suit the needs of the visually impaired even if such terminals provide only limited functions. To facilitate the implementation of such a system, most of its components would preferably be situated in the environment. Terminals would feature the most simplified structure possible and would be easy to operate, such that the system would not place an undue burden on users.

Based on the foregoing principles, we have proposed a voice-guidance system[9] that incorporates infrared communication and simple terminals based on the technologies used in CoBIT[8]. This system is constructed based on a simple mechanism: the sound waves of voice guidance are converted into an infrared signal and transmitted; a photoelectric cell in the receiver terminal then receives the infrared signal and reconstitutes the sound wave, and the user listens to the voice guidance through an earbud directly connected to the photoelectric cell (Fig. 4). This system allows for the use of low-cost transmitters and terminals, and in addition, it places only a small burden on the user: the terminals are small, and the user can begin receiving voice signals simply by turning on the receiver.

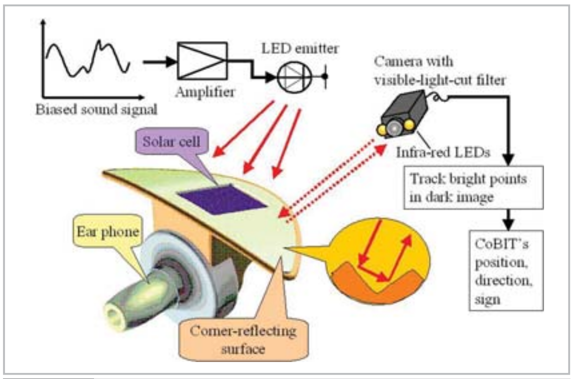

Fig.4 Mechanism of infrared communication based on CoBIT

We performed an application experiment on this system, with the cooperation of a number of visually impaired subjects. The results show that the directionality and locality of infrared communication is effective in communicating directions intuitively and in providing information according to the position of the user. More generally, these results indicate that a system constructed with these simple devices also offers significant potential as a voice-guidance system.

On the other hand, the nature of the infrared signal described above limits the coverage area, and it often proved difficult to capture the source of the infrared signal, leading the user to wander off the route. Accordingly, we need to incorporate measures that will ensure that the terminals will reliably receive the transmitted infrared signal.

3.2 Voice guidance system combining infrared communication and AM radio broadcasting

To solve the problem described in the previous sub-section, we considered incorporating a means of setting limits to the user's range of mobility. Specifically, we came up with a technique of indicating to the terminal where it should look for the infrared transmitter, rendering it easier for the terminal to receive the infrared signal. The technique we devised to restrict the range of mobility involved the adoption of a weak AM radio wave transmitter. AM broadcasting is a type of non-directional communication; the use of weak radio waves enables broadcasting within a range of several meters. Based on these characteristics, we developed a technique enabling the terminal to receive the infrared signal more easily by first notifying the terminal of the approximate position of the infrared transmitter.

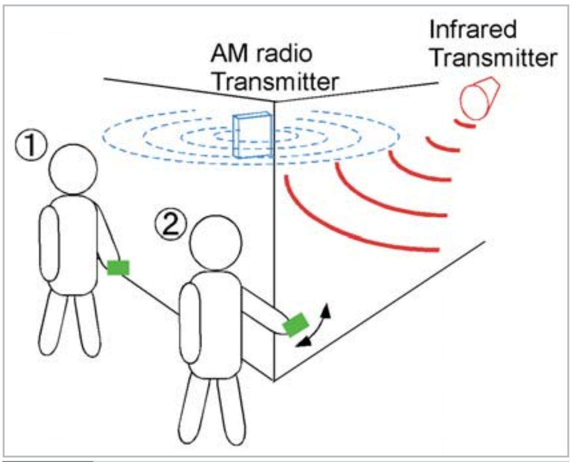

Fig.5 Schematic diagram of mobility support by AM-CoBIT

Figure 5 shows a schematic diagram of the use of the voice guidance system combining infrared communication and AM radio broadcasting (referred to as AM-CoBIT below). At a position requiring caution, such as in front of an intersection, the AM radio broadcasting function presents the voice message, "Approaching an intersection", letting the user know of the intersection in advance [(1) in Fig. 5]. The user locates the intersection based on the volume of the received voice message from the AM transmitter; the user then moves the terminal to the right and left to allow the terminal to detect the infrared signal; voice guidance is then issued to the user via the infrared transmitter (on the left in the diagram) regarding location and direction, indicating "(the destination) is this way". [(2) in Fig. 5]. The foregoing are the assumed procedures under the system.

This type of method has already been implemented and is in operation within a system to assist pedestrians at intersections (the Pedestrian Information and Communication System, or PICS), which combines FM radio broadcasting and infrared communication[10].

However, we adopted the AM broadcasting due to certain advantages over FM broadcasting — for example, AM features less detuning noise, and transmission volume is more distance-dependent, allowing for the position of the transmitter to be roughly estimated based on the loudness of the signal.

AM-CoBIT consists of an infrared transmitter, a weak radio wave transmitter for AM radio, and a receiving terminal. The infrared transmitter does not use the conventional method of converting a voice into an infrared signal; instead, the device transmits an infrared signal based on the pulse-width modulation (PWM) method, which adjusts the duty ratio of rectangular pulses at a specific frequency according to the amplitude of the voice signal input to the transmitter. For AM radio broadcasting, we use the weak AM radio-wave transmitter adopted by Tekuteku Radio[11]. The radio transmitter has a recording IC circuit, a modulating circuit, and a loop antenna built-in, and transmits weak radio waves that satisfy the conditions for stations not requiring radio licenses as specified in the Radio Law. More specifically, the transmitter has sufficient power to allow the terminal to receive radio waves within an area of approximately 5 m from the transmitter. For the carrier wave frequency of the AM radio broadcasting, we chose 1,620 kHz — the band used to broadcast road traffic information — to avoid interference with general broadcasting.

On the other hand, the receiving terminal consists of a commercially available AM radio receiver and a circuit that switches between the AM radio signal and infrared signal for output to the earbud. Normally, the circuit outputs voice from the AM radio, but when the terminal detects an infrared signal containing certain frequency components on the receiving surface of the terminal, the circuit switches automatically to infrared signal input and outputs voice converted from the infrared communication to the earbud.

Here, although the system uses infrared communication based on the PWM method, conventional amplitude modulation is employed as the signal is input to the CoBIT earbud, allowing the user to hear voice through CoBIT. Broadcasting based on a weak AM radio transmitter can also be heard using an ordinary AM radio, so the system can also provide useful voice guidance to the physically unimpaired, depending on the voice content. This feature will allow the system to be more readily incorporated into a new information infrastructure.

3.3 Application experiment

3.3.1 Overview of experiment

We performed an experiment to evaluate the usefulness of the system described in **3.2** by asking a number of visually impaired subjects to participate in its use. Fifteen visually impaired people (fully blind) assisted us in this experiment.

As shown in Fig. 6, the experiment was performed indoors in hotels and shops. The AM radio transmitter transmitted announcements regarding the location — for example, “Approaching an intersection” or “Approaching a toilet”. The infrared transmitter transmitted information related to specific location and direction — for example, “The elevator is this way”, “The entrance is this way”, and “The toilet is this way”. The examinees traveled dis-

Fig.6 *Scenes of AM-CoBIT application experiment*

tances of approximately 30 to 100 m.

Here, the AM radio transmitter used the voice of a man while the infrared transmitter used the voice of a woman, so the user could intuitively distinguish the voice from AM radio transmission from the voice coming from infrared communication switching. We explained this structure to the subjects before the experiment.

3.3.2 Experimental results

This experiment focused on how users use the system to proceed at a branching point such as an intersection. Accordingly, we observed the users acting as we had expected: finding the approximate position of the intersection using the voice from the AM radio; proceeding little by little as they moved the terminal left and right to search for the infrared transmitter; and finding the correct direction in which to proceed. In interviews after the experiment many examinees commented that they used the AM radio indications for a rough sense of what was near and then paid closer attention to the voice from the infrared signal to find their way.

We also found that the voice of the AM radio functions as a preliminary alert. When we performed an experiment with a voice guidance system using only infrared communication, many examinees commented that they wanted to be alerted before arriving at intersections or facilities. One of the examinees said, “I knew from the AM radio voice that I was near a toilet, but it was only when the voice changed that I was able to locate the actual entrance to the toilet”. This example shows that the AM radio functions as an initial alert.

On the other hand, several problems also became clear, first among them related to the use of the terminal. For example, when the terminal received the infrared signal, the user could not identify the direction of the infrared transmitter and proceeded in the wrong direction. This problem may be overcome as users become accustomed to walking with the devices, but we believe that in any case it is necessary to improve the terminal design or other elements to ensure that the user can intuitively identify the direction of the signal by simply holding the terminal. Some problems were also noted with respect to the communication method, including the sound quality of the received voice. It often took time for the subject to understand the content of the voice guidance. Improvements in the means of communication method are thus required.

We also need to reconsider the content of the voice guidance and the positioning of the transmitters. As indicated above with respect to the alert function, we can conclude that the change in the voice guidance itself functioned as an alert. To ensure that we make the most effective use of this function, we need to position the transmitters so that this change in voice guidance takes place smoothly and reliably.

3.3.3 Discussion

The results discussed above show that this system, combining directional voice guidance information based on infrared transmission with non-directional broadcasting based on AM transmission, is effective in indicating direction of movement and in confirming location. By providing positional information several meters before arriving at an object using the AM transmitter, the system serves to provide useful alerts. Recent advances in information technology have reduced the total number of transmitters required to perform these functions; as a result, the system does not require as many terminals as we feared would be necessary when studying a system using only infrared signals.

Although several areas of improvement remain, this system seems to offer many advantages — for example, it can be implemented at low cost and it can also be used as a voice guidance system for the physically unimpaired through adjustment of voice content. The terminal is also simple, an advantage in encouraging adoption of the system.

4 Conclusions

Though able to walk, many people face difficulties in moving about due to impaired vision or hearing. To improve the quality of life of the disabled under these circumstances, it is important to develop a system that provides information that will counteract their difficulties in acquiring information on the outer world. In this context, the functions of a user-carried mobile terminal in the RCT system discussed above have been developed with the aims of minimizing the burdens of technical implementation and keeping the initial costs of introducing the system to a minimum. We believe that these systems are ready for practical application.

As a final note, we would like to express our sincere gratitude to the hearing and visually impaired subjects who cooperated in our experiments and all others who helped us in managing the experiments.

References

1 I. Yairi, K. Kayama, and S. Igi, "Robotic Communication Terminals as a Mobility Support System for Elderly and Disabled People (5)", Journal of Japanese Society for Artificial Intelligence Vol.20 No.1, pp.82-89, 2005. (in Japanese)

2 H. Matsuo, S. Lu, S. Igi, K.Imagawa, Y. Tanaka, and Y. Nagashima, "A Japanese Sign Language recognition based on global movement featrures and localized characteristics in images captured using stereo camera", The Transaction of Human Interface Society Vol.3 No.3, pp.135-144, 2001. (in Japanese)

3 S. Igi, R. Watanabe, and S. Lu, "Synthesis and Editing Tool for Japanese Sign Language Animation", The Journal of the Institute of Electronics, Information and Communication Engineers, Vol.J84-D-1 No.6, pp.987-995, 2001. (in Japanese)

4 S. Lu, H. Sakato, T. Uezono, and S. Igi, "Synthesizing Japanese Sign Language using Intuitive Motion Primitive", Proc. of 8th HCI International '99 Vol.8, pp.441-445, 1999.

5 T. Uezono, H. Sakato, S. Lu, and S. Igi, "Development and Experiment of Synthesizing System for Japanese Sign Language Animation Based on Motion-Primitives and Mouth-Shapes", IEICE technical report HIP96-46, pp.49-56, 1999. (in Japanese)

6 T. Shimada et al., "Information Guidance System for the Handicapped in a Railway Station", IEICE technical report WIT2001-6, pp.29-34, 2001. (in Japanese)

7 Thomas STROTHOTTE, et al, "MoBIC: user needs and preliminary design for a mobility aid for blind and elderly travelers", 2nd TIDE congress, 1995.

8 T. Nishimura et al., "A Compact Battery-less Information Terminal System for Location-based Interactive Information Support", Transactions of Information Processing Society of Japan Vol.44 No.11, pp.2659-2669, 2003. (in Japanese)

9 S. Oyama et al., "Walking Support by Infrared Audio Communication and Compact Low-battery Terminals", Journal of the Japan Society of Infrared Science and Technology Vol.13 No.2, pp.33-42, 2004. (in Japanese)

10 T. Tajima, et al., "Pedestrian Information and communication systems for visually impaired persons", Proc. of CSUN's 16th Annual International Conference "Technology and Persons with Disabilities", 2001.

11 The audio guidance system with the feeble electric wave "Tekuteku Radio" website, http://tekuteku-radio.com/

***OYAMA Shin'ya**, Ph.D.*

Assistant Professor, Department of Computer Engineering, Hakodate National College of Technology (Former: Associate Researcher, Human-Comuputer Intelligent Interaction Group, Information and Network Systems Department)

Human Interface

***NISHIMURA Takuichi**, Ph.D.*

Group Leader, Real-World Based Interaction Group, Information Technology Research Institute, National Institute of Advanced Industrial Science and Technology (Former: Guest Researcher, Human-Comuputer Intelligent Interaction Group, Information and Network Systems Department)

Retrieval and Recognize of Time-Sequence Pattern, Event Space Information Support

***IGI Seiji**, Ph.D.*

Managing Director, International Alliance Division, Research Promotion Department (Former: Director, Keihanna Human Info-Communication Research Center)

Welfare Information Technology, Human Interface

***YAIRI, Ikuko Eguchi**, Ph.D.*

Research Manager, Universal City Group, Knowledge Creating Communication Research Center (Former: Senior Researcher, Human-Comuputer Intelligent Interaction Group, Keihanna Human Info-Communication Research Center, Information and Network Systems Department)

Welfare Information Technology, Human Interface