

A Study of Universal Accessibility for People in Japan with Communicative and Perceptual Disorders

Tohru Ifukube

Research Institute for Electronic Science

Hokkaido University,

Sapporo 060-0812, Japan

ifukube@sense.es.hokudai.ac.jp

Abstract

The author has developed one basic research approach for universal accessibility over a period of 28 years. As reviewed in this paper, he and his co-researchers have designed several devices for universal accessibility as well as 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 human-centred computer interfaces such as virtual reality, automatic speech recognition and speech syntheses. Moreover, these newly developed computer interface technologies have led to the improvement in the design of models for developing universal accessibility devices.

1. Introduction

We have been carrying out the universal accessibility studies and have designed several aid devices which are now or shall soon be put into practical use for people in Japan with communicative and perceptual disorders [1]. Moreover, we have obtained many findings concerning both the mechanisms of perception and concept formations in the human brain. In this paper, we will refer to our research regarding the aid devices and explain how that research has been related to new information technologies such as virtual reality, speech recognition and speech syntheses. Lastly, we will introduce our research approach to developing human-centred interfaces.

2. SOME UNIVERSAL DESIGN DEVICES for COMMUNICATIVE and PERCEPTUAL DISORDERS

2.1. Digital Hearing Aid for the Elderly Hearing Impaired

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 1.

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 sensori-neural hearing-impaired [2]. This hearing aid has been manufactured by Hitachi Ltd. It was priced at \$400 (US).

We have also proposed a digital processing method to modify tone contrast for the hearing impaired, especially for Chinese people. Each Mandarin monosyllable has 4 tones which determine the meaning of each Chinese character. Figure 2 shows 4 tones of /ma/ in Chinese. The tone contrast is defined as the greatest difference in frequencies between peaks and valleys of pitch curves in monosyllable utterances. Under quiet and noisy backgrounds, modified Mandarin tone words were presented to hearing-impaired Chinese listeners with moderate to severe sensori-neural

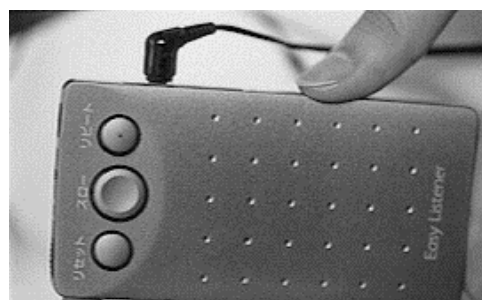


Fig.1: Digital hearing aid with a speech rate conversion function

hearing loss. The listeners were asked to identify four alternative monosyllable words which were distinguishable by tones 1, 2, 3 and 4 respectively.

Employing this method, we found that modified speech with enhanced tone contrast yielded moderate gains in the percentage of correct identification of the tones when compared to unmodified speech tones with only compression amplification. The finding therefore offers support to the assertion that a hearing aid with tone modifications is indeed effective for the hearing-impaired, especially for the elderly [3].

We are planning to design a new digital hearing aid which has both speech rate conversion and tone modification functions.

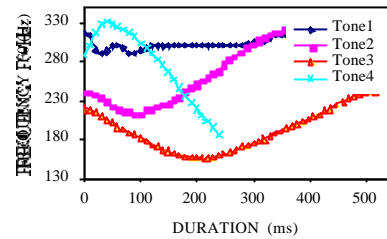


Fig.2: 4 tones of /ma/ in Chinese

2.2. Electric Artificial Larynx for People with Speech Production Disorders

Various methods for vocal rehabilitation have been applied to patients who have undergone a laryngectomy, 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.

An electric artificial larynx is one of the artificial larynges for people, especially for the elderly who are hard to master other forms of substitute speech. However, these prostheses make it difficult to produce a natural sounding voice. From the analysis of the mynah's vocalization mechanism, we have found that the mynah can clearly imitate phonetic information such as intonation and pitch fluctuation [4]. 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 patients who have undergone laryngectomies to control intonation by using their respiration.

The device consists of three parts, as shown in figure 3. 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. The third part is an electromechanical vibrator that can be attached to the neck.

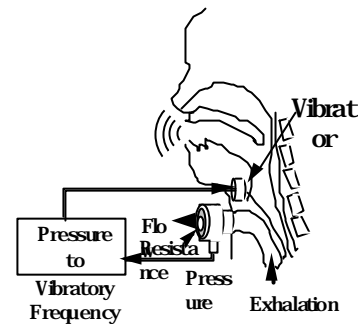


Fig.3: Block-diagram of the electric larynx

This type of electric artificial larynx has just been manufactured by a company in Japan. It has been proven that intonation is very important in making an artificial larynx voice sound natural. Pitch patterns obtained from a subject after one day of training became clearly similar to the patterns produced from a normal subject [5]. Up to now, about 1000 devices have already been sold in Japan. Figure 4 shows a manufactured electrical larynx. It is priced at \$700 (US).



Fig 4 : Manufactured electric larynx

From a questionnaire given to 100 users, it was found that about 30% of them reported using both the pitch control function and the singing mode using melodies stored onto the computer memory. The remaining of users reported that they have not tried to control the pitch because it seemed to be difficult to master the pitch control function without any instruction [6].

From the above report, we have improved the human interface which are necessary to facilitate its use in addition to instruction needed to teach users the pitch control function.

Furthermore, the technology for the electric larynx has been applied to design a speech synthesis for a screen reader as mentioned next.

2.3. A GUI Screen Reader for the Visually Impaired in Japan

In Japan, word processors for visually impaired people were first developed in the early 1980s. A few years later, screen readers for MS-DOS® were also developed. Recently, however, operating systems based on graphical user interfaces (GUIs) have spread rapidly in personal computer environments. Consequently, they are inherently difficult for visually impaired people to use since they are hard to use of visual information.

With this problem in mind, we have proposed screen reader software as a supporting tool for GUI operation by visually impaired people in Japan. The proposed software transmits the information on the screen to visually impaired users through synthesized speech. It does not require the use of pointing devices; all computer operations can be done from the keyboard. Our screen reader builds an environment in which visually impaired people can use

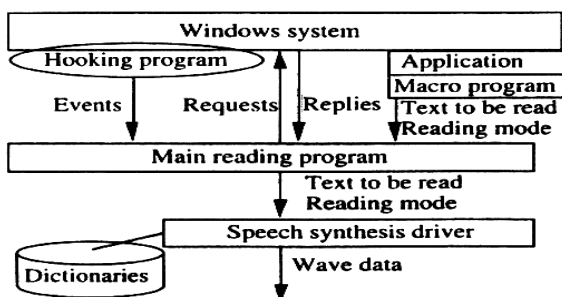


Fig. 5 : Structure of the screen reader

applications like word processors on their own in Microsoft Windows. In view of the above, we developed a prototype screen reader in 1996 called the 95-reader. Our screen reader for Windows is composed of several pieces of software [7].

To improve the softwares for visually impaired people, an MS-DOS front-end processor with speech output was revised to work in the Windows environment. This program not only converts Japanese syllables (called kana) into Chinese characters (called kanji), but also outputs them into verbal expressions that are understandable as speech. From the evaluation tests by 52 visually impaired users, we are confident that the keyboard commands

of our screen reader are sufficiently easy to operate. However, some subjects reported that it would be better for the user to be able to change the speed at will.

Based on the evaluation results, the prototype screen reader was revised. For example, reading mode was added as a parameter to be sent from the main reading program to the speech synthesis driver, thereby the speed, the pitch, the length of pauses at punctuation, the accent and the speaker (male, female, or computer) can be chosen by the user. Figure 5 shows the structure of the revised version.

The revised screen reader was put into practical use in 1997 and has been widely spread throughout Japan. It is priced at \$ 300 (US). The screen reader has been improved to make it compatible with Windows 98 and Windows 2000. Figure 6 shows a package of the 95 reader which is compatible with Windows 2000. Almost 80% of visually impaired users in Japan are employing our screen reader.



Fig.6 Package case of the screen reader

In the future, it will be necessary for visually impaired people to have more Windows applications with vocalizing functions. Furthermore, if they can utilize a LAN application, including the use of Internet mail, visually impaired workers will be able to send and receive documents with people all over the world, thereby improving their working environment.

2.4 Tactile Display for Hearing and/or Visual Aids

Through fundamental research on auditory and tactile information processing, we have developed a tactile voice coder called a tactile vocoder. This device was manufactured for hearing-impaired people in Japan about 25 years ago[8]. The tactile vocoder is a device which produces sound spectral patterns that are analysed in 16 frequency components. When an index fingertip touches a piezo-electric vibrator array, the device makes it possible to discriminate the first and the second formants of vowels as well as some consonants such as fricatives, semi-vowels and nasals. As well as being put to practical use in Japan, studies on its evaluation were conducted at some institutes for the hearing impaired [9]. Figure 7 shows a prototype of a portable tactile vocoder. It was priced at more than \$ 7000 (US).

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 vocoder) after one week of training. Although it was put into practical use, it has not been widely spread because of its high price and large size.

Recently, with the help of the Japanese national government, we have designed a new tactile display for communication aids [10]. This display provides new sensory information for the hearing and/or visually impaired. The device includes a tactile

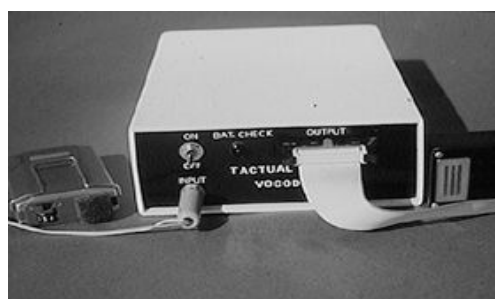


Fig.7 :Prototype of portable tactile vocoder

stimulator matrix consisting of 16 rows and 4 columns and a small-sized wearable computer.

The tactile display is made up of 64 vibrating pins fixed onto each piezoelectric element at 1 or 2 mm intervals. The tactile display is 2 cm long, 1 cm wide and 1 cm high so that it can be attached to a finger tip as shown in figure 8. The vibration intensity can be changed for each vibrator from 0 to 40 dB. and its vibratory frequency can be set to around 200 Hz. The vibrating pins stimulate the surface of a fingertip in order to display a speech spectrum of a complete sentence.

By slowly displaying the speech spectrum, the subject is able to catch the spectrum's movement. Moreover, by displaying the vibration pattern from right to left on the fingertip like an electric scoreboard, the subject can more clearly recognize the rapidly changing spectral pattern. By adopting such slowed down and sweeping methods, the average identification rate for Japanese 7 monosyllables reaches 70 % after one week of training [11].

The tactile converter consists of a computer, with drivers for the vibrators and rechargeable batteries, that can be worn on a person's arm much like a bracelet. The device, including the tactile display and converter, will be offered at the low price of around \$500 (US).

This technology can be used together with a hearing aid, sign language, lip-reading and cochlear implants. Speech signals obtained from cellular phones and e-mail voice messages can also be converted into vibratory patterns. Those patterns may likewise be used to obtain feedback of a subject's own speech signals for pronunciation training. They could also enable a person to catch alarm sounds while walking on a street or even to detect emergency information broadcasts from radios.

In order to display the images drawn on a paper or computer screen, a pixel of the images is detected by a CCD sensor attached to the bottom of the stimulator matrix. Each density of the images is then converted into a vibratory intensity corresponding to each pixel. By scanning a whole image, the subjects will be able to construct the corresponding image inside the brain.

This tactile display can further be used to catch an outline or shade of images, pictures and graphs presented on a computer display with a GUI. Moreover, in conjunction with the screen reader mentioned above, the visually impaired will be able to communicate with computers by themselves. They will also be capable of detecting the brightness of their surroundings by pointing the CCD camera in the direction they want to go.

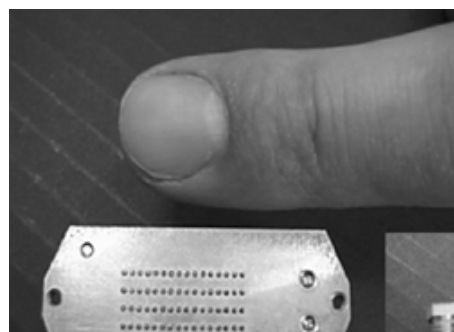


Fig.8: New vibrator array

2.5. Voice Recognition System with an HMD for the Hearing Impaired



Fig.9 :Prototype voice typewriter

Our tactile vocoder was applied to a voice typewriter and an 8 channel cochlear implant for those with acquired deafness about 20 years ago, since it is especially difficult for such hearing impaired to learn lip-reading and sign language. As to the cochlear implant, the author conducted its basic research at Stanford University about for one year. Although many basic findings were obtained from psychophysical experiments with help of a volunteer to whom our electrode array was implanted [12], [13], we could not continue the cochlear implant research in Japan because of lack of cooperation in both medical field and Japanese companies. The technology of our

cochlear implant was applied to an implantable tinnitus suppressor for people with a ringing ear called tinnitus [14]. On the other hand, the voice typewriter was successfully designed using a

micro-computer at that time. This device can convert monosyllabic voice sounds such as /a/, /ka/, /sa/, /ta/, /na/, /ha/ into Japanese letters almost in real time and show them on a display [15].

In order to better conceptualize this process, some explanation regarding the Japanese language may be helpful. 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. These characteristics make it easy to design a voice typewriter for the Japanese language. Figure 9 shows the prototype voice typewriter. Its recognition rate was 96% and its response time was about 0.2 seconds. It was priced at more than \$ 10000 (US). The voice typewriter has been applied to the input device of Japanese word processors.

Recently, this technology was applied to design a portable speech recognition system with a see-through Head Mounted Display (HMD) through which the hearing impaired would be able to read spoken sentences as a series of letters together with lip-reading, facial expressions, and gestures of a speaker.

However, modern speech recognition technology is not perfect; a fact which results in frequent recognition errors in places where humans would often be able to guess the correct meaning of sentences through context. By making use of this guessing ability, the present speech technology could prove useful as a communication aid for the hearing impaired; especially for the acquired deaf.

Actually, from our experimental results, it was found that the subjects can accurately guess the meaning of the sentences if about half of the speech has been correctly recognized. We have also investigated whether or not non-verbal information such as speaker's lip-movement and facial expressions would help to understand sentences which contain incorrect words. For this experiment, both the letters and facial expressions were displayed through a see-through HMD. The results suggest that combining both incomplete verbal and non-verbal information is indeed significant in facilitating comprehension [16].



Fig.10: New hearing assistive device with a speech recognition and a see-through HMD

In view of this finding, we have proposed a new hearing assistive device which utilizes a portable speech recognition system combined with a see-through eyeglass display for the acquired hearing impaired, as shown in figure 10.

2.6. Obstacle Detection for the Visual Impaired

Using microprocessor and ultrasonic devices, a new model for a mobility aid has been developed to enable visually impaired people to perceive their surroundings. 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 11.

Ultrasonic reflections from obstacles are picked up by a two-channel receiver. The frequency of the emitted ultrasound is swept from 70 to 40kHz within 1ms, 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.

With this device, a visually impaired person could detect a 2mm diameter wire at a distance of about 1 meter. We have also found that this device enables the visually impaired to discriminate between several obstacles at the same time without the use of any virtual images. This mobility aid, modeled after a bats' echo-location system, is very effective at detecting small obstacles placed in front of the head [17].

However, we have not put the device into practical use because sounds presented through the earphones of the device seem to interfere with auditory information obtained from one's surroundings. Actually, most visually impaired people can detect obstacles without these devices by using an ability known as the "obstacle sense".

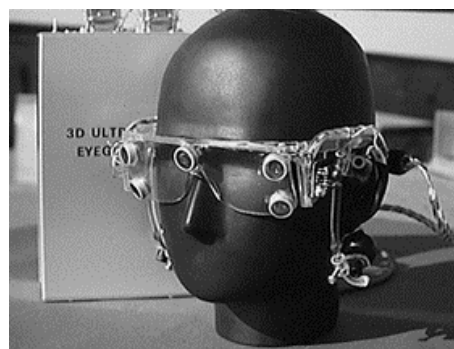


Fig.11 : Prototype Ultrasonic eyeglasses

There are two steps involved in the obstacle sense: first, perception and afterwards, appraisal. We have been investigating the mechanism of the obstacle sense based on psychophysical experiments using visually impaired students. From this research, we have ascertained that the reason why they can detect the obstacle is due to their ability to discriminate the tiny difference in sound field between themselves and an obstacle compared with when there is no obstacle present. Furthermore, we were able to make the blind 'hear virtual obstacles' by controlling the sound field produced from a speaker array which is shown in figure 12. This study has also been related to virtual reality research. In the near future, we intend to design a mobility aid device which will likewise make use of the ability of the obstacle sense.

2.7. Virtual Reality Studies for Perceptual Disorders

Based on the above techniques and findings, we have constructed a virtual reality (VR) system which consists of a head-mounted display, a speaker array, a rotational chair, and a soundproof room (figure 13). At the same time, our laboratory has been conducting a research project on a mixed reality (MR) system since 1997 in order to construct a kind of augmented reality. This is being done in cooperation with the Japanese 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. Before applying the MR technologies to aids for perceptual disorders, we evaluated how the MR stimulation influenced the human body; in particular, the visual function, the autonomic nervous system, and the sense of equilibrium.



Fig.12: Speaker array for displaying virtual obstacles

In order to investigate these influences, the Sapporo research branch was constructed. At the centre, we have three rooms: a control room, a display room and a bio-medical measurement room.

Based on the findings obtained from the evaluation experiments, we will determine how the MR technologies can be applied as an aid for perceptual disorders without any harmful effects [18].

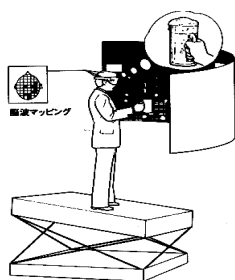


Fig.13 Arch-screen and a motion base with two force plates equipped inside the MR system laboratory

Up to now, we have proved that MR technologies are useful in designing an auditory aid with both a speech recognition device and a see-through HMD, as mentioned at section 2.3. We have also found that an image conversion technique using an HMD with a video camera is useful for increasing the visual area for spatial perception disorders. Furthermore, from basic experimental results, we have suggested that a moving sound image added to a moving visual image is helpful for rehabilitation of the sense of balance, especially for the balance sense of the elderly.

3. CONCLUSION

The basic findings concerning human sensory functions and universal accessibility devices will be useful for designing human-centred interfaces. In the near future, the newly developed technologies of human-centred interfaces will be applied to the design of better models of assistive technologies for people with perceptual and communicative disorders. This is our research approach of universal accessibility (UA) as schematically shown in figure 14 [19].

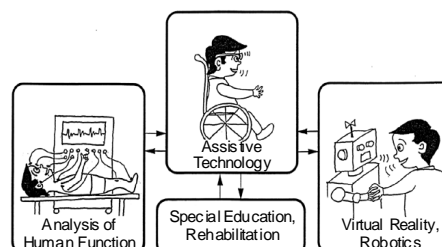


Fig.14 :One basic research approach for UA

Acknowledgement

The author thanks Shuichi Ino, Norihiro Uemi, Chikamune Wada, Norishige Nagai, Yoshito Nejime, Tetuya Watanabe and graduate students involving in this study for their help in running the experiments and designing the devices. Funding was supported by Grant-in-Aid from the Ministry of Education and Science (No. 10305031).

References

- [1] IFUKUBE, T., Sound-based Assistive Technology for the Disabled, in: T. Ifukube, Corona Publishing Co., (250 pages. in Japanese) (1997)
- [2] NEJIME, Y., ARITSUKA, T., IFUKUBE, T., MATSUSHIMA, J., A portable digital speech rate converter for hearing impairment, in: IEEE Trans., Rehabilitation Eng., 4(2) 73-83 (1996)
- [3] LU, J., UEMI, N., LI, G., IFUKUBE, T., Tone Enhancement in Mandarin Speech for Listeners with Hearing Impairment. IEICE Transactions on Information and Systems (2001, in press)
- [4] AOKI, N., IFUKUBE, T., 1/f shimmer and jitter in normal sustained vowels and their application to speech synthesis, in: J.Acoust.Soc.Am (in press) (1998)
- [5] UEMI, N., IFUKUBE, T., Design of a new electrical larynx having a pitch control function, in: IEEE Inter. Workshop on Robot and Human Communication, 198-203 (1994)
- [6] IFUKUBE, T., UEMI, N., A New Electrical Larynx with Pitch Control Function. Proceedings of the 2nd East Asian Conference on Phonosurgery, 1-5(1999)
- [7] WATANABE, T., OKADA, S., IFUKUBE, T., Development of a GUI screen reader for blind persons, Systems and Computers in Japan. 29, 13, 18-27(1998)
- [8] IFUKUBE, T., A cued tactile vocoder, in: Computers in Aiding the Disabled, J. Raviv (ed), North-Holland Publishing Co., 197-215 (1982)
- [9] 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)
- [10] IFUKUBE, T., WADA, C., INO, S., SHOJI, H., TSUNASHIMA, H., On a project for practical use of a hearing assist device using the tactile sense. Technical Report of IECE, WIT00-14, 13-18(2000)
- [11] 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)
- [12] 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)
- [13] 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)
- [14] MATSUSHIMA, J., MIYOSHI, S., IFUKUBE, T., Surgical method for implanted tinnitus suppresser, in: Inter. Tinnitus Journal, 2(1) 21-25 (1996)
- [15] IFUKUBE, T., Design of a Voice Typewriter, in: T. Ifukube, CQ Publishing Co., (195 pages, in Japanese), (1984)

[16] KOSAKAI, A., NARA, H., INO, S., YOSHIDA, J., IFUKUBE, T., A speech recognition assist device for the hearing impaired which presents both in complete verbal information and non-verbal information. Technical Report of IECE, 13, 7-12 (2000)

[17] IFUKUBE, T., SASAKI, T., PENG, C., A blind mobility aid modeled after echolocation of bats, in: IEEE Trans., BME 38(5) 461-465 (1991)

[18] IFUKUBE, T., A guideline for the design of 3D-displays based on physiological parameters for use in an MR environment, Proceedings of Inter, Symp. on Mixed Reality, Special Session: Final Report of the MR Project, 116-124 (2001)

[19] IFUKUBE, T., One basic research approach of assistive engineering in Japan, KAJZER, J., TANAKA, E., YAMADA, H., (Eds) Human Biomechanics and Injury Prevention, Springer-Verlag Tokyo, 191-198 (2000)