Tactile Vision - Development of a Wearable Prosthesis for the Blind

Martin Gonda
Intelligent Robotics Research Centre
Dept. Electrical & Computer Systems Engineering
Monash University, Wellington Road
Clayton, Victoria 3168, Australia
E-mail: Martin.Gonda@eng.monash.edu.au

Ray Jarvis
Intelligent Robotics Research Centre
Dept. Electrical & Computer Systems Engineering
Monash University, Wellington Road
Clayton, Victoria 3168, Australia
E-mail: Ray.Jarvis@eng.monash.edu.au

Abstract

This paper describes the development of a wearable, stereo-camera based prosthetic device, allowing visually impaired persons to perceive a 'depth-image' of the scene ahead, conveyed using vibrating actuator cues. Preliminary tests indicate that such a device may be a viable and preferable alternative to existing aids including those based on ultra-sonic sensors and audible cues, while other alternatives such as retinal and cortical implants, are still only in the early stages of development.

1 Introduction

Given our reliance on vision to perform most day to day tasks, blindness or a significant loss of eyesight would certainly be one of the most feared disabilities, particularly as we get older.

Despite many medical advances, people who were born blind or have lost their sight due to illness or accident will generally remain so for the remainder of their lives. Nevertheless, recent electronic technology developments and bio-engineering research into artificial retinas [Gross et al, 1998; Stieglitz et al, 1997] and cortical brain implants [Dobelle, 2000] may help restore at least partial vision to some patients in the future. It will still be a number of years (or even decades) however, before safe, effective, as well as commercially viable implants are available, and it is also quite possible that many patients will not be suited for the procedure due to the specific nature of their impairment.

Apart from restoring actual eyesight, many other tools and methods currently exist to aid the visually impaired. Some utilise ultra-sonic sensors, while others employ infra-red laser range-finders, and often audible cues are generated to warn the person about possible obstacles [Borenstein & Ulrich, 1997; Kaczmarek, 1995; Kay 1984; Meijer 1992]. However, rather than providing a three dimensional view of the scene ahead, the actual information often tends to be one-dimensional in the sense that only the nearest object tends to register in the sensor's field of view. Additionally, there is a possibility that the audible feedback may distract the person from hearing the traffic, while sound cues are often of little use for those that are also hearing impaired.

This paper proposes a wearable prosthetic device, depicted in Figure 1, which should overcome many of these shortcomings. It consists of an array of vibrating actuators arranged in a grid and worn on a person's back, conveying the information about the 3D environment in the scene ahead in the form of an 'image' of tactile cues.

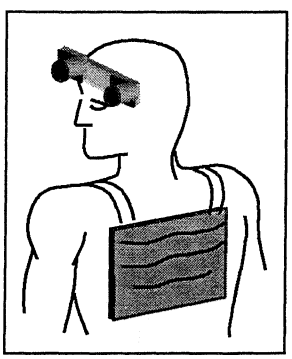

Figure 1. Tactile Vision device

The front-end of the system is comprised of a stereo camera board, along with a PC and stereo image matching software, which can generate range related disparity maps at near frame rate. This data in turn drives the actuator grid by pulsing the vibrating motors that make up each tactile actuator.

Similar wearable haptic displays for the blind have been implemented and studied over the past two or three decades [Jansson, 1983; Kaczmarek et al, 1991; Kay 1984; Saunders, 1983; Tan, 1997], but most haven't been utilised for conveying three-dimensional information, and have instead focused on portraying intensity images, using either vibro-tactile or electro-tactile cues. While devices such as these have had some success, and have been adapted for other parts of the body, predominantly the fingers, but also the tongue [Bach-y-Rita at al, 1998; Kaczmarek et al, 1997], they have generally been used only for conveying simple patterns and shapes (eg. letters and words for reading) and for educational purposes. Displaying intensity images of the environment using these devices has only been successful in contrived environments, where the background is generally uncluttered so as not to confuse the user.

Our method effectively removes the background since it conveys depth rather than intesity information, and therefore does not suffer from the above problem. It also does not need any intrusive and possibly dangerous surgical procedures as requierd by retinal and cortical implants, and at the same time it should prove to be more informative to the visually impaired person than distance only devices. It is also likely to be more intuitive to use and understand than devices which provide only audible feedback, since describing or representing a scene using sound is quite a difficult task. Furthermore, no disruption of useful environmental sounds (which are particularly precious to the blind) is involved.

2 Hardware

The hardware employed for this project (Fig.2) is made up of readily available and inexpensive parts. Firstly, we use a stereo camera board developed by SRI [Konolige, 1999], which is made up of two miniature CCD camera modules with interchangeable lenses, and also contains interlacing circuitry for combining the scanline data from the two cameras into a single video stream. This eliminates the need for dual video capture cards, although separate video streams are also supported by the camera board.

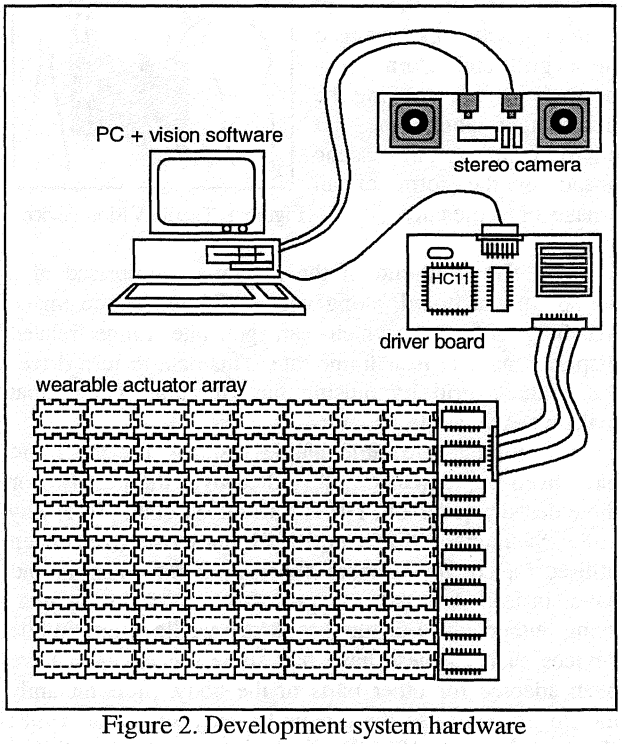

Figure 2. Development system hardware

The stereo image processing is currently performed on a compact and 'wearable' (if only just, at present) PentiumIII-based system, which also includes a Pxc200 (Brooktree chipset) video capture board from *Imagenation*. The test software and the libraries for performing stereo image matching, were provided by SRI, along with the stereo camera board.

The field of view for the stereo cameras is between 80 and 130 degrees, depending on the focal length of the lenses used. This field of view is further reduced by 20 to 40 degrees after the disparity matching has been performed, depending on the separation of the cameras, and on some of the variables used during matching process (eg. the size of the matching window, and the number of allowable disparities).

2.1 Actuator Grid

The actuator grid, depicted in Figure 2, constitutes an array of 64 small vibrating DC motors arranged in an 8x8 grid. A microcontroller board is used to drive the motors by means of a PWM signal, while also serving as an interface between the computer and the actuator array. Each row of the array also comprises a small addressable latching driver IC chip, allowing each row to be driven in sequence, and reducing the number of interconnecting wires. The driver board, which may eventually be integrated and mounted with the motor array, communicates with the computer system via a standard RS232 serial interface, requiring only a 3-wire connection.

Each individual actuator consists of a small DC motor, mounted in a flexible moulded silicon compartment as shown in Figure 3. Each motor has on off-balance weight attached to the shaft, thus producing vibrations as it rotates. Such motors are commonly used in portable paging equipment as well as in vibrating batteries of mobile phones and can be sourced from surplus electronics stores, with the weight already attached.

The individual silicon casings are designed to interlock by means of hollow conduits, allowing the wiring for each motor to be neatly hidden, while at the same time holding the array together.

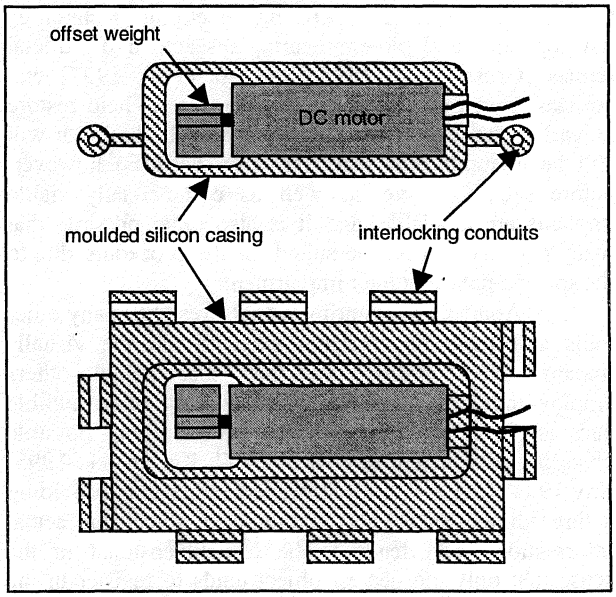

Figure 3. Cross-section view (side and top) of vibrating actuator

While part of the actuator grid has been fabricated, including the silicon encased vibrating motors (Fig 4), to aid in the feasibility testing of this conceptual system, a somewhat crude subset of the actuator array has been constructed by glueing several motors to an elastic band, which can then be placed against any suitable part of the body.

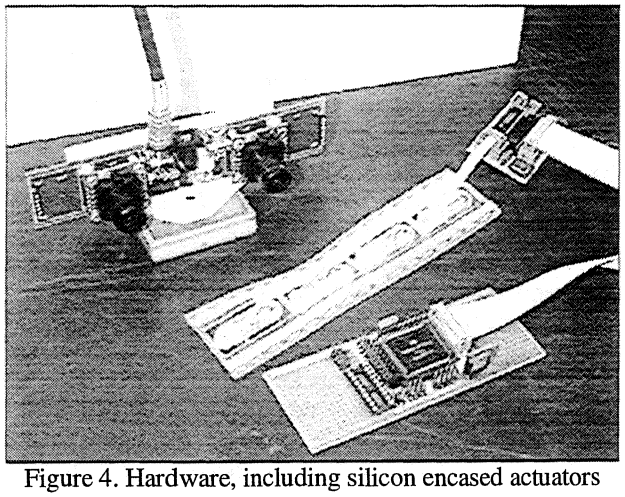

Figure 4. Hardware, including silicon encased actuators

3 Software

The libraries provided by SRI for performing the stereo matching tasks are used in our system to generate a disparity map image (Fig 5). This can then be easily converted into actual distances to individual objects in the scene, given that the camera parameters (eg. separation) are known. However, this extra step can be bypassed, since we are only concerned about conveying relative depth information to the visually impaired user by means of varying the strength and/or the frequency of the vibrations in the different parts of the 'image'. The user should then be able to readily learn the association between different vibrating sensations and the actual depths.

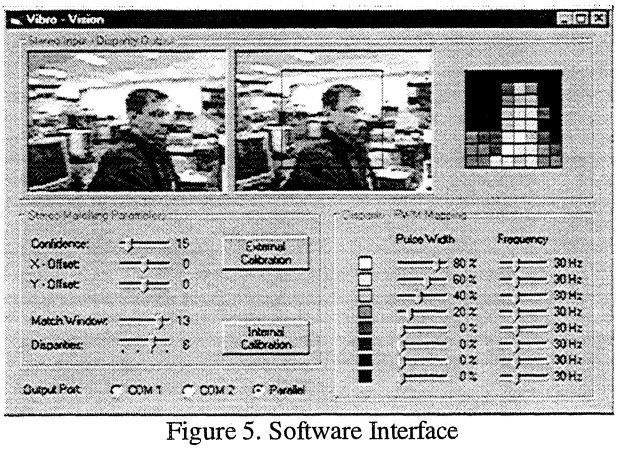

Figure 5. Software Interface

Since the resolution of the actuator array is relatively coarse, a method of reducing the resolution of the actual disparity image is required. This can be accomplished by a number of methods, including applying a low-pass filter to the image, or perhaps just taking the maximum value (corresponding to the shortest distance) for each block of pixels.

The second method ensures that small objects of close proximity to the blind person, which are likely to present the most immediate obstacle, are not filtered out. However, since the stereo matching algorithm is not 100% accurate, spurious matches may often indicate objects nearby which are not actually there. Therefore a combination of the above two methods is envisaged, and correlation with previous frames might also be employed in the final system.

4 Testing and Results

The main focus of the preliminary tests has been identifying whether the proposed prosthetic device would be able to convey sufficient resolution of depth information using the available vibrating motors, and whether the user would be able to differentiate between the different levels of vibration. Similar tests have been carried out by Kaczmarek et al [1991], among others.

The first proposed location of the body for wearing our actuator array was a person's back. It presents a large area of skin which can be in contact with the device, and allows the actuators to be sufficiently spread apart, so as to allow distinguishing between them. It also doesn't present any problems with regards to reducing the person's mobility.

Given a set of four vibrating motors distributed approximately 4cm apart, the author, as well as other subjects, were able to differentiate between at least 4 distinct disparity levels (2 bits of resolution per actuator). It is believed that a trained person should be able to increase their perception to at least 8 different levels (3 bits). However, the spacing of the actuators within the grid will probably have to remain at approximately 4cm, as it is already difficult to spatially distinguish between adjoining vibrating motors. This corresponds to studies by Kaczmarek et al [1991, 1995], which indicate spacing of approximately 3-5cm. This is attributable to the density of the nerve cells on a person's back, which is much lower than that of the fingertips or the tongue, for which our device is obviously unsuitable.

Another interesting effect was observed when the frequency of the PWM signal which drives the DC motors was reduced to only a few Hertz, allowing the person to feel the individual pulses. The perceived sensation was totally different to simply varying the intensity of the vibrations at higher PWM rates, suggesting that the resolution per actuator could be increased further. Alternately, activating the motors in such a manner may be useful in conveying information about potentially dangerous obstacles which are may be very close to the subject or which are moving (such as vehicles or people).

5 Suggestions and Conclusions

Early tests indicate, that a wearable vibro-tactile vision device, which conveys 3-dimensional information, such as the one described in this paper, could become a very useful and viable aid for many visually impaired patients.

The proposed device is currently still in the development stages, and therefore not very portable due to the bulky processing hardware. However, the research team at SRI have constructed much more compact DSP- as well as FPGA-based hardware for performing the same processing tasks of the stereo video image matching and generating a disparity map. It is envisaged that should the wearable prosthetic device become viable and receive further funding, a smaller unit, weighing perhaps a kilo in total, can be developed.

Additional improvements might constitute adding ultra-sonic sensors for disambiguating closest objects in noisy disparity images, or for environments where the images are not sufficiently bright, such as in dark rooms or at night.

Other forms of stimulating the touch senses, such as electro-tactile devices, or smaller Braille-type actuators may also be considered.

References

[Bach-y-Rita, 1983] P. Bach-y-Rita, "Tactile Vision Substitution: Past and Future", *International Journal of Neruroscience*, Vol. 19, 1983, pp. 29-36.

[Bach-y-Rita et al, 1998] P. Bach-y-Rita, K.A. Kaczmarek, M.E. Tyler, J. Garcia-Lara, "Form Perception with a 49-Point Electrotactile Stimulus Array on the Tongue", *Journal of Rehabilitation Research and Development*, Vol. 35, No. 4, Oct 1998, pp. 427-430

[Borenstein & Ulrich, 1997] J. Borenstein and I. Ulrich, "The GuideCane - A Computerised Travel Aid for the Active Guidance of Blind Pedestrians", *Proc. of IEEE International Conference on Robotics and Automation*, Apr 1997, pp. 1283-1288

[Craig, 1983] J. C. Craig, "Some Factors Affecting Tactile Pattern Recognition", *International Journal of Neruroscience*, Vol. 19, 1983, pp. 47-58.

[Dobelle, 2000] W. H. Dobelle, "Artificial Vision for the Blind by Connecting a Television camera to the Visual Cortex", ASAIO *(American Society of Artificial Internal Organs) Journal 2000*, Vol. 46, pp. 3-9

[Gross et al, 1998] M. Gross, R. Buss, T. Alder, R. Heinzelmann, D. Kalinowski, D. Jager, "Artificial Vision: An Applications for Short Distance Free Space Optical Interconnection", *Optics In Computing*, Jun 1998, Brugge, Belgium.

[Jansson, 1983] G. Jansson, "Tactile Guidance of Movement", *International Journal of Neruroscience*, Vol. 19, 1983, pp. 37-46.

[Kaczmarek et al, 1991] K. A. Kaczmarek, J.G. Webster, P. Bach-y-Rita, W.J. Tompkins, "Electrotactile and Vibrotactile Displays for sensory Substitution Systems", *IEEE Transactions on Biomedical Engineering*, Vol. 38, No. 1, Jan 1991, pp. 1-16

[Kaczmarek, 1995] K. A. Kaczmarek, "Sensory Augmentation and Substitution", *CRC Handbook of Biomedical Engineering*, pp. 2100-2109, CRC Press, 1995

[Kaczmarek et al, 1997] K. Kaczmarek, M.E. Tyler, P. Bach-y-Rita, "Pattern Identification on a Fingertip-Scanned Electrotactile Display", *Proceedings - 19th Int. Conf. - IEEE/EMBS*, 1997, pp. 1694-1696.

[Kay, 1984] L. Kay, "Electronic Aids fo Blind Persons: An Interdisciplinary Subject", *IEE Proceedings*, Vol. 131, Pt. A, No. 7, Sep 1984, pp. 559-576.

[Konolige, 1999] K. Konolige, "SRI Small Vision System: User's Manual, Software Ver1.3", May 1999, SRI International, (http://www.ai.sri.com/~konolige)

[Meijer, 1992] P.B.L. Meijer, "Experimental System for Auditory Image Representation", *IEEE Transactions on Biomedical Engineering*, Vol. 39, No. 2, Feb 1992, pp. 112-121.

[Saunders, 1983] F.A. Saunders, "Information Transmission Across the Skin: High Resolution Tactile Sensory Aids for the Deaf and the Blind", *International Journal of Neruroscience*, Vol. 19, 1983, pp. 21-28.

[Stieglitz et al, 1997] T. Stieglitz, H. Beutel, R. Keller, C. Blau, J-U. Meyer, "Development of Flexible Stimulation Devices for a Retina Implant System", *Proceedings - 19th Int. Conf. - IEEE/EMBS*, 1997, pp. 2307-2310.

[Tan et al, 1997], Hong Z. Tan, Ifung Lu, A. Pentland, "The Chair as a Novel Haptic User Interface", *Proceedings of the Workshop on Perceptual User Interfaces*, Oct 1997, pp. 56-57