Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

A Navigation Aid for the Blind Using Tactile-Visual Sensory Substitution

Lise A. Johnson and Charles M. Higgins*

Abstract— **The objective of this study is to improve the quality of life for the visually impaired by restoring their ability to self-navigate. In this paper we describe a compact, wearable device that converts visual information into a tactile signal. This device, constructed entirely from commercially available parts, enables the user to perceive distant objects via a different sensory modality. Preliminary data suggest that this device is useful for object avoidance in simple environments.**

I. INTRODUCTION

The visual sense provides a rich and complex set of information about the surrounding environment, and in particular it informs an organism about the positions and properties of objects in the world. Generally humans rely very heavily on visual input to perform a host of important functions including facial and object recognition, depth estimation, navigation and object avoidance. When deprived of this information, however, a person may learn to compensate by relying on input from other sensory modalities. Visual sensory substitution can take a variety of forms and may be mediated by a device. The traditional white cane is perhaps the simplest example of a visual sensory substitution device; information about distant objects that could have been provided visually is instead collected by the tactile receptors of the hand. Obviously, this method is very limited compared to vision. The temporal and spatial resolution of a cane is poor and the potential for error is correspondingly high. However, such problems are tied to the method of information acquisition and are not implicit limitations of the tactile sense.

Visual substitution studies have focused on both the tactile and auditory senses. Using both modalities, subjects have demonstrated the ability to localize and identify objects in the visual field, and even to determine their depth using looming, perspective or motion parallax [1]. Visual-auditory substitution, however, taxes a sensory modality that is already extensively used for communication and localization. Therefore, in this study we have chosen to focus on visual-tactile substitution. The skin is an ideal location for visual substitution for several reasons. First, the area of the skin is very large and the great majority of it is minimally used during navigation. In addition, like the retina, the skin is capable of representing information in two dimensions and integrating signals over time [2], which means that visual patterns and tactual patterns are functionally interchangeable [3]. Furthermore, a number of visual illusions have also been demonstrated for the tactile sense, indicating that perception of spatio-temporal sequences is not exclusively determined at the sensory level but is in fact a feature of central nervous system processing [4], [5].

Tactile-visual sensory substitution has been investigated by many groups since the 1960's [2], [6], [7], [8], [9], [10], [11], [12]. The Tactile Vision Substitution (TVS) system [13] is an important example of a tactile-visual substitution device. In this system visual information was captured with a video camera and delivered via a tactile array to the skin of the back, abdomen or thigh. With the TVS system, users were able to perform complex "eye"-hand coordination and facial recognition tasks. Furthermore, as with vision, these users learned to project the stimulus origin to a location distant from the site of stimulation [14]. However, the TVS system proved to be inadequate for use in normal (cluttered) environments. Kaczmarek and Bach-y-Rita [3] postulated that the reason for this might be that the abundance of detail provided by the system overwhelms the information that is important for navigation in a complex environment. Attempts have been made to get around this problem by automatically identifying and classifying approaching objects and then providing the user with cues corresponding to orientation, relative position, direction and range, in a coded form, either audibly or through a tactile display [15]. However, in addition to the fact that automatic recognition and classification is a difficult problem for computer vision, visual information presented in a coded form, even if it is accurate, is not intuitively useful for the subject [16]. Spatial information is more readily interpreted when presented spatially.

In the design of this navigation aid we have focused on circumventing the resolution issues that have hindered visual-tactile sensory substitution devices in the past. By concentrating on object avoidance, rather than object recognition, we are able to reduce the amount of information to be conveyed to manageable proportions. Our strategy is to extract the salient features of the visual input in a pre-processing stage, and then provide this information in a spatially relevant way via a sparse tactile array. Other groups have used similar tactics to develop navigational devices. The drawbacks inherent in these approaches have to do with the mode of presentation. Velazquez et al. [8] supply the visual signal to a two-dimensional array of moveable pins that the user holds with one hand and scans

---

*Lise A. Johnson is a Ph.D. student in the Biomedical Engineering Program at the University of Arizona, Tucson. email: `laj1@email.arizona.edu`

Charles M. Higgins is with the Department of Electrical and Computer Engineering/Biomedical Engineering Program at the University of Arizona, Tucson. email: `higgins@ece.arizona.edu`

1-4244-0033-3/06/$20.00 ©2006 IEEE.

with the other. This effectively occupies both hands, which is undesirable because it restricts other activities. Areibi and Zelek [12] opted to condense all of the information into five binary tactors which were placed on the back of the user's hand. The quantity and quality of information that can be conveyed by such a system is very limited. In both cases, the location of the display on one of the hands is sub-optimal because the brain often references directions with respect to the relatively stable torso and not to the extremities [17]. The substantial differences in our design have to do with the site and method of stimulation and the quality of the information that is supplied.

II. DEVICE DESIGN

Before building this navigation aid, it was necessary to define the aspects of the visual scene that constitute the most important features for navigation. This is not a simple task, because while humans use a wide range of information for navigation (depth, form, color, motion, etc.) it is not clear what information can best be interpreted tactually. It is clear, however, that the user must be able to identify an object's presence and its position in space in order to avoid colliding with it. For this project we have designated the three-dimensional position of an object as the feature of greatest value for navigation.

Visual images, whether projected onto a living retina or recorded by a camera, are two-dimensional projections of a three-dimensional world. When three-dimensional information is required it must be reconstructed from these projections. There is more than one way to accomplish this, but in this study we have chosen to implement a computational stereo vision algorithm.

When choosing a site on the body to apply this information to, it is important to consider both the sensory qualities of the skin and the comfort of the subjects. We have chosen the skin of the abdomen as the stimulation site. Although the spatial resolution of the sensory receptors in this area is much less than other parts of the body (e.g. the fingertips, face, and tongue [18]) the abdomen has the advantage of being relatively flat, large, and easily accessible. Most importantly, the abdominal skin is not heavily taxed for other sensory uses and therefore stimulation at this site does not interfere with other functions that are important for navigation [2].

III. THE PROTOTYPE DEVICE

Fig. 1. Photograph of the navigation aid. a) The tactor belt, containing 14 vibrating motors spaced laterally. b) The camera belt, with two webcams arranged to converge approximately 10 feet from the user. c) The full device from the side, as it is meant to be worn. Visual processing is performed by a laptop computer inside the backpack.

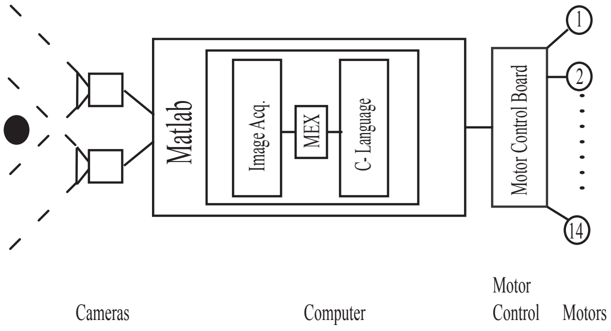

Fig. 2. Schematic representation of the prototype. At each time step, a frame is acquired by each camera. These frames are processed by a computational stereo algorithm, the output of which is non-linearly converted into a vibrotactile stimulus. One or more of the motors vibrates when an object is detected in their assigned portion of the visual field. A higher frequency of vibration indicates a closer object.

Our prototype device is built entirely from off-the-shelf hardware components including two webcams, a laptop computer, fourteen hobby servos modified to run DC motors, a servo control board, and fourteen miniature vibrating motors attached to a flexible belt. The fourteen vibrating motors are fully insulated by plastic tubing so that no electrical components are exposed to the user. These tubes are sewn around the perimeter of a fabric belt at a spacing of approximately 2 cm (see Fig. 1a). The spacing of the tactors was chosen to agree with the acuity of the torso for vibrotactile stimuli as measured by van Erp [19], who reported a uniform acuity between 2 and 3 cm except on the body midline where acuity was approximately 1 cm. The two webcams are mounted on a rigid platform and attached to a belt that is worn around the user's waist. The cameras are vertically registered but are not calibrated. The angle between the cameras is fixed, with the point of convergence approximately ten feet in front of the user (see Fig. 1b). The camera belt is worn over the tactor belt. The motor control boards and the battery are housed in a plastic box that is worn by the user in a backpack along with the laptop computer that processes the visual input. The motor wires and camera cables run from the user's waist into the backpack (see Fig. 1c).

Fig. 2 shows schematically the processing steps involved in converting the visual (camera) input into a tactile stimulus. A computational stereo algorithm assigns a three-dimensional position to any object falling in the overlapping region of the cameras' fields of view (the effective field of

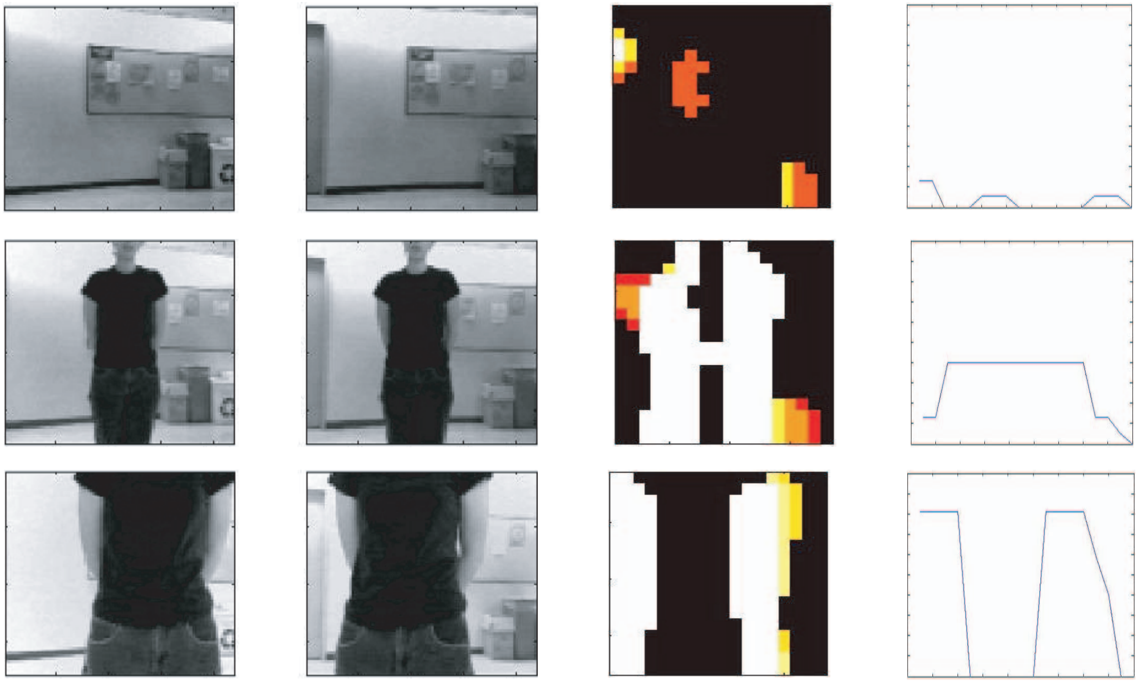

Fig. 3. An example of the device output. Each row shows (from left to right) an image from the left and right cameras, the disparity map, and the corresponding signals sent to the tactor belt. These are frames taken of the same scene as a person approaches. Note the non-linear increase in activity as the person comes closer. In the last row the tactor signal is bifurcated because the depth algorithm uses edges and texture, features which are absent in the center of the object.

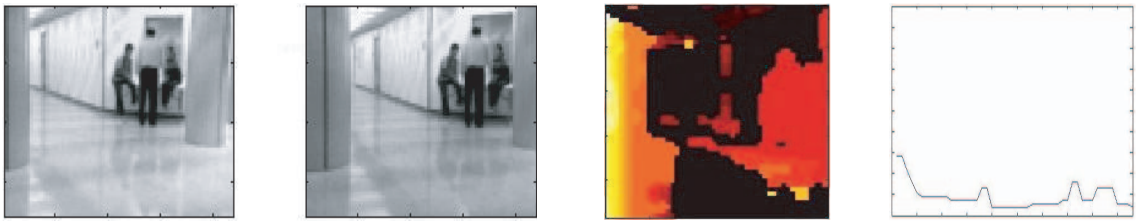

Fig. 4. An example of the device output for a more complicated scene. From left to right this figure shows the two stereo frames of a hallway captured by the cameras, the disparity map, and the corresponding signals sent to the tactor belt.

view). The stereo algorithm is written in the C programming language, compiled by the Matlab compiler (MEX), and run in the Matlab interpreter. The images are acquired in Matlab with the Image Acquisition Toolbox. The motors are activated by Matlab and controlled by a servo control board. The disparity depth map is two-dimensionally median filtered to remove random noise and then divided into fourteen vertical, overlapping sections. This two dimensional map is condensed into one dimension by assigning the value of the nearest object in each slice to the entire section of the field. Each tactor is assigned a particular section and is activated when an object falls in this portion of the visual field. The nearer the object is, the higher the frequency of vibration of the motor. Thus, if a nearby object were detected on the far left of the effective field of view, the left-most tactor would respond with a relatively strong vibratory signal. A non-linearity in the algorithm causes the response to increase dramatically as an object approaches. The algorithm is tuned to respond most strongly to objects within a range that is useful for walking navigation. Objects that are too distant or too close are ignored. The device has been completely built and integrated, and data has been collected in real indoor environments. The prototype processes up to 10 frames per second, which is sufficiently fast for normal walking speeds.

IV. PRELIMINARY DATA

Fig. 3 shows a series of frames captured by the cameras, the depth map that is computed by the stereo algorithm, and the command signal that is sent to the tactor array. When objects are far enough away from the cameras the motor command signal is small. As an object/person approaches, the magnitude of the command signal, and thus the frequency of vibration of the motors, increases rapidly. In the same way Fig. 4 shows the output of the system for a more complicated indoor scene such as a user might encounter in everyday life. The pole on the left appears in the tactor signal as a moderately strong stimulus, but not strong enough to indicate an imminent collision. The group of people and the second pole on the right are detected by the system, but do not produce a strong stimulus as they are relatively far away. There is no information in the output that would allow the user to distinguish between different types of obstacles (e.g. to tell the difference between a pole and a person) and thus this device is strictly useful for object detection and not object recognition.

V. DISCUSSION

The results presented in this paper mark the beginning of our efforts to build a compact, wearable navigation aid that allows the visually impaired to negotiate everyday environments in real time. Using entirely off-the-shelf components we have developed a device that gathers visual inputs, extracts the information relevant for navigation (where the nearest objects are in three-dimensional space) and subsequently translates it into a tactile signal. In this preliminary implementation of the device the visual data is acquired by inexpensive webcams and the tactile signal is provided by an array of vibrating motors. In the future it is our intention to replace these cumbersome and power-hungry components with VLSI vision chips [20] and an electro-tactile array [21]. The VLSI chips will be custom hardware implementations of the same algorithms that are currently implemented in software. This allows them to be smaller, faster and more efficient. Likewise, electro-tactile stimulation will allow the device to be smaller and more efficient but will not require any change in tactor configuration.

In addition to these hardware improvements, we are also investigating different depth algorithms and different configurations of the tactor array. We are interested in testing the limits of human ability to understand the environment with sparse sensory input. To this end we are using our prototype device as a testing tool to investigate what type and what quality of information is minimally necessary for navigation. We are also investigating whether there is a point of information saturation beyond which perceptual improvements are minimal. These questions are interesting from the perspective of device engineering, but no less interesting with respect to brain plasticity.

REFERENCES

[1] P. Bach-y-Rita and S. Kercel, "Sensory substitution and the human-machine interface," *Trends in Cognitive Sciences*, vol. 7, pp. 541–546, 2003.

[2] F. A. Geldard, "Some neglected possibilities of communication," *Science*, vol. 131, pp. 1583–1588, 1960.

[3] K. Kaczmarek and P. Bach-y-Rita, "Tactile displays," in *Virtual Environments and Advanced Interface Design*, W. Barfield and T. A. Furness, Eds. Oxford, New York, 1995, pp. 541–546.

[4] F. A. Geldard and C. E. Sherrick, "Space, time and touch," *Scientific American*, vol. 255, pp. 90–95, 1986.

[5] E. C. Lechelt, "Temporal and intensitive factors in comparative judgments of tactile space," *Percpetual and Motor Skills*, vol. 48, pp. 179–185, 1979.

[6] F. A. Geldard and C. E. Sherrick, "Multiple cutaneous stimulation: The discrimination of vibratory patterns," *The Journal of the Acoustical Society of America*, vol. 37, pp. 797–801, 1965.

[7] E. C. Lechelt, "Sensory substitution systems for the sensorily impaired: the case for the use of tactile-vibratory stimulation," *Perceptual and Motor Skills*, vol. 62, pp. 356–358, 1986.

[8] R. Valazquez, E. E. Pissaloux, J. C. Guinot, and F. Maingreaud, "Walking using touch: Design and preliminary prototype of a non-invasive ETA for the visually impaired," *Proceedings of the 2005 IEEE in Medicine and Biology 27th Annual Conference*, 2005.

[9] B. W. White, "Perceptual findings with the vision-substitution system," *IEEE Transactions on Man-Machine Systems*, vol. 11, pp. 54–58, 1970.

[10] J. C. Bliss, M. H. Katchner, C. H. Rogers, and R. P. Shepard, "Optical-to-tactile image conversion for the blind," *IEEE Transactions on Man-Machine Systems*, vol. 11, 1970.

[11] P. Bach-y-Rita, "Tactile sensory substitution studies," *Annals of the New York Academy of Sciences*, vol. 1013, pp. 83–91, 2004.

[12] A. Shawki and Z. J., "A smart reconfigurable visual system for the blind," *Proceedings of the Tunisian-German Conference on: Smart Systems and Devices*, 2001.

[13] P. Bach-y-Rita, C. C. Collins, F. A. Saunders, B. W. White, and L. Scadden, "Vision substitution by tactile image projection," *Nature*, vol. 221, pp. 963–964, 1969.

[14] P. Bach-y-Rita, *Brain Mechanisms in Sensory Substitution*. New York, Academic Press, 1972.

[15] C. C. Collins and M. F. Deering, "A microcomputer based blind mobility aid," *Proceedings of the IEEE Frontiers of Engineering and Computing Healthcare*, pp. 52–55, 1984.

[16] B. White, F. Saunders, L. Scadden, P. Bach-y-Rita, and C. Collins, "Seeing with the skin," *Perception and Psychophysics*, vol. 7, pp. 23–27, 1970.

[17] H. O. Karnath, P. Schenkel, and F. B., "Trunk orientation as the determining factor of the 'contralateral' deficit in the neglect syndrome and as the physical anchor of the internal representation of body orientation in space," *Brain*, vol. 114, pp. 1997–2004, 1991.

[18] S. Weinstein, "Intensive and extensive aspects of tactile sensitivity as a function of body part, sex, and laterality," in *The Skin Senses*, D. R. Kenshalo, Ed. Charles C. Thomas, 1968, pp. 195–222.

[19] J. B. F. van Erp, "Vibrotactile spatial acuity on the torso: effets of location and timing parameters," *Proceedings of the First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, 2005.

[20] C. M. Higgins and V. Pant, "Biomimetic VLSI sensor for visual tracking of small moving targets," *IEEE Transactions on Circuits and Systems*, vol. 51, pp. 2384–2394, 2004.

[21] H. M. Seifert and A. Fuglevand, "Restoration of movement using functional electrical stimulation and Bayes' theorem," *Journal of Neuroscience*, vol. 22, pp. 9465–9474, 2002.