

A Substitute Vision System for Providing 3D Perception and GPS Navigation via Electro-Tactile Stimulation

Simon Meers and Koren Ward
School of IT and Computer Science
University of Wollongong
Wollongong, NSW, Australia, 2522
meers@uow.edu.au, koren@uow.edu.au

Abstract

This paper describes a novel approach for enabling the blind to achieve obstacle avoidance and navigation in outdoor environments with the aid of visual sensors, GPS and electro-tactile stimulation. The electro-neural vision system (ENVS) works by extracting a depth map from stereo cameras by measuring the disparity between the stereo images. This range data is then delivered to the fingers via electro-neural stimulation to indicate to the user the range of objects being viewed by the cameras. To perceive the location of obstacles and the 3D structure of the environment the user imagines that the hands are held in the direction viewed by the cameras, with fingers extended, and the amount of stimulation felt by each finger indicates the range of objects in the direction pointed at by each finger. Also, the relative location of significant landmarks is determined using GPS and stored GPS coordinates and delivered to the fingers via encoded pulses when the landmarks are in the field of view of the stereo cameras. This intuitive means of perceiving the 3D structure of the environment and the location of landmarks in real time effectively enables the user to navigate the environment without use of the eyes or other blind aids. Experimental results are provided demonstrating the potential that this form of 3D environment perception has at enabling the user to achieve localisation, obstacle avoidance and navigation without using the eyes.

Keywords: substitute vision, TENS, electro-tactile, electro-neural vision, stereo cameras, disparity, GPS

1 Introduction

It is estimated there are 37 million blind people worldwide [1]. Consequently, blindness not only represents a severe cognitive handicap, but also a significant burden on the global community. To address this problem, we have been experimenting with electro-tactile user interfaces, stereo video cameras and GPS for providing the user with useful 3D perception of the environment and landmarks without using the eyes. Our vision system (shown in Figure 1) works by extracting depth information from the stereo cameras and delivering this information to the fingers via electro-neural stimulation. To interpret the range data, the user simply imagines that the hands held with fingers extended in the direction viewed by the cameras. The amount of electro-neural stimulation felt by each finger indicates the distance to objects in the direction of each of the fingers.

Furthermore, GPS and stored GPS coordinates are used to indicate to the user the relative location of significant landmarks in the user's vicinity. This is achieved by encoding significant landmarks as a sequence of pulses and delivering the appropriate pulse sequence to the fingers periodically when the corresponding landmark is estimated to be within the field of view of the stereo cameras. The distance to

perceived landmarks is also encoded by the intensity of the pulse sequence.



Figure 1. The Electro-Neural Vision System

By having environmental depth information and the location of landmarks delivered continuously to the user in a form that is easy to interpret, the user is able to realise the 3D profile of their surroundings and the location of known landmarks in the environment by surveying the environment with the cameras. This form of 3D perception can enable the user to navigate the environment, recognise the user's location and perceive the size and movement of objects within the surrounding environment without using the eyes.

In Section 2 of this paper we provide a brief review of previous work done on artificial or substitute vision systems and GPS aids for the blind. In Section 3 we provide implementation details of the Electro Neural Vision System (ENVS) with respect to 3D perception and landmark recognition. In Section 4 we provide the results of experiments we have conducted with the ENVS to demonstrate its possible use for assisting the blind to navigate outdoor environments. Finally, we provide concluding remarks and a brief description of further work to be done.

2 Background

Bionic vision in the form of artificial silicon retinas or external cameras that stimulate the retina, optic nerve or visual cortex via tiny implanted electrodes are currently under development (see [2], [3] & [4]). Currently, the only commercially available artificial vision implant is the Dobbelle Implant [5]. This device provides visual perception in the form of points of light that bear little resemblance to the surrounding environment. Even if more successful results are achieved with implants in the not so distant future, many blind people may not benefit from implants due to their high cost and the expertise required to surgically implant the device. Some forms of blindness (eg. brain or optic nerve damage) may also be unsuitable for implants.

In addition to bionic vision implants, a number of wearable devices have been developed for providing the blind with some means of sensing or visualizing the environment. For example, Meijer's vOICe [6] compresses a camera image into a coarse 2D array of greyscale values and delivers this information to the ears as a sequence of sounds with varying frequency.

Considerable work on sonar mobility aids for the blind has been done by Kay [7]. Kay's systems deliver frequency modulated sounds, which represent an object's distance by the pitch of the generated sound and the object's surface texture by the timbre of the sound delivered to the headphones. However, to an inexperienced user, these combined sounds can be confusing and difficult to interpret. Also, the sonar beam from these systems is very specular in that it can be reflected off many surfaces or absorbed resulting in uncertain perception. However, Kay's sonar blind aids can help to identify landmarks by resolving some object features (i.e. resolution, texture, parallax) and can facilitate some degree of an object's classification to experienced users.

A further drawback of auditory substitute vision systems is that they can diminish a blind person's capacity to hear sounds in the environment, (eg voices, traffic, walking, etc). Consequently, these devices are not widely used in public places because they can reduce a blind person's auditory perception of the environment and could potentially cause harm

or injury by reducing a blind person's capacity to detect impending danger from sounds.

Electro-tactile displays for interpreting the shape of images on a computer screen with the fingers, tongue or abdomen have been developed by Kaczmarek *et al* [8]. These displays work by mapping black and white pixels to a matrix of closely spaced pulsated electrodes which can be felt by the fingers. Although these electro-tactile displays can give a blind user the capacity to recognise the shape of certain objects, like black alphabetic characters on a white background, they do not provide the user with any useful 3D perception of the environment which is needed for environment navigation, localization, landmark recognition and obstacle avoidance.

In addition to sensing the surrounding environment, it is of considerable benefit if a navigational aid can provide information regarding the absolute position of the user. This can be achieved by using GPS technology. Currently, several devices are available which perform this task for the blind. (e.g. Drishti [9], BrailleNote GPS [10], Trekker [11], and others [12], [13]). The processing of GPS data to provide non-visual guidance is already well explored. However, as with audio substitute vision systems, by occupying the sense of hearing this can diminish a blind person's capacity to hear important environmental sounds.

Our ENVS [14] is significant because it provides a useful intuitive means of perceiving the 3D structure of the environment which does not impede a blind person's capacity to hear sounds in the environment. In this paper we describe how the relative location of landmarks in the environment can also be perceived via the ENVS. This makes it possible for a user to negotiate obstacles in the environment while navigating the environment using perceived landmarks. The user can also manually enter landmarks in the ENVS while negotiating the environment so that the user can easily remember the location of visited location. In the following section, we provide a brief description of the ENVS hardware equipped with GPS and its operation.

3 ENVS Implementation

The basic concept of the ENVS is shown in Figure 2 and is comprised of a stereo video camera headset fitted with a digital compass for obtaining video and direction information from the environment, a laptop computer equipped with a GPS unit for processing the video and direction data from the headset and obtaining the GPS location of the user, a Transcutaneous Electro-Neural Stimulation (TENS) unit for converting the output from the computer into appropriate electrical pulses that can be felt via the skin, and special gloves fitted with electrodes for delivering the electrical pulses to the fingers.

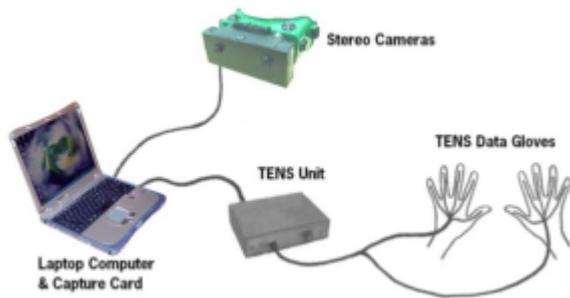


Figure 2. The basic concept of the ENVS

3.1 Obtaining 3D Perception from the ENVS

The ENVS extracts depth information from the environment by using the laptop computer to obtain a disparity depth map of the immediate environment from the head mounted stereo cameras. This is then converted into electrical pulses by the TENS unit that stimulates nerves in the skin via electrodes located in the TENS data gloves that can be seen in Figure 3. To achieve electrical conductivity between the electrodes and skin, a small amount of conductive gel is applied to the fingers prior to fitting the gloves. For our test purposes, the stereo camera headset is designed to completely block out the users eyes to simulate blindness.



Figure 3. The ENVS setup

The key to obtaining useful environmental information from the electro-neural data gloves lies in representing the range data delivered to the fingers in an intuitive manner. To interpret this information the user imagines his or her hands are positioned in front of the abdomen with fingers extended. The amount of stimulation felt by each finger is directly proportional to the distance of objects in the direction pointed by each finger.

To enable the stereo disparity algorithm parameters and the TENS output waveform to be altered for experimental purposes we provided the ENVS with the control panel shown in Figure 4. This was also designed to monitor the image data coming from the cameras and the signals being delivered to the fingers via the TENS unit.

Figure 4 shows a typical screen grab of the ENVS's control panel while in operation. The top-left image

shows a typical environment image obtained from one of the cameras in the stereo camera headset. The corresponding disparity depth map, derived from both cameras, can be seen in the top-right image (i.e. lighter pixels have a closer range than darker pixels). Also, the ten disparity map sample regions, used to obtain the ten range readings delivered to the fingers, can be seen spread horizontally across the centre of the disparity map image. These regions are also adjustable via the control panel.

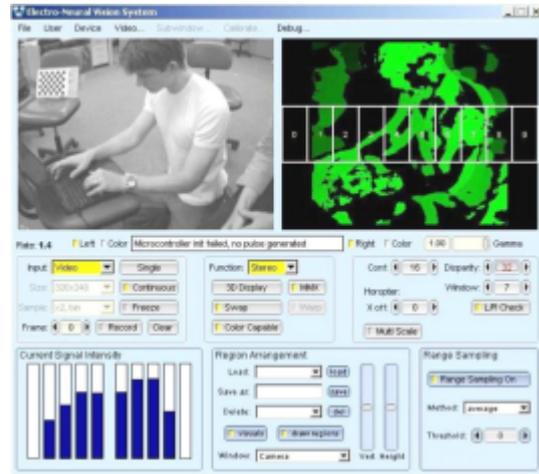


Figure 4. The ENVS control panel.

To calculate the amount of stimulation delivered to each finger, the minimum depth of each of the ten sample regions is taken. The bar graph, at the bottom-left of Figure 4, shows the actual amount of stimulation delivered to each finger. Using a 1000 MHz Pentium computer we were able to achieve a frame rate of 15 frames per second which proved more than adequate for our experiments.

3.1.1 Extracting Depth from Stereo Video

The ENVS works by using the principle of stereo disparity [see [15)]. Just as our eyes capture two slightly different images and our brain combines them with a sense of depth, the stereo cameras in the ENVS captures two images and the laptop computer computes a depth map by estimating the disparity between the two images. However, unlike binocular vision on humans and animals, which have independently moveable eye balls, typical stereo vision systems use parallel mounted video cameras positioned at a set distance from each other.

3.1.2 Limitations

To calculate the disparity between image pixels the stereo disparity algorithm requires features to be detected in the stereo images. Consequently, featureless surfaces can pose a problem for the disparity algorithm due to a lack of identifiable features. For example, Figure 5 illustrates this problem with a disparity map of a whiteboard. As the whiteboard surface has no identifiable features on it,

the disparity of this surface and its range cannot be calculated. To make the ENVS user aware of this, the ENVS maintains a slight signal if a region contains only distant features and no signal at all if the disparity cannot be calculated due to a lack of features in a region. With further work we expect to overcome this deficiency by also incorporating IR range sensors into the ENVS headset.

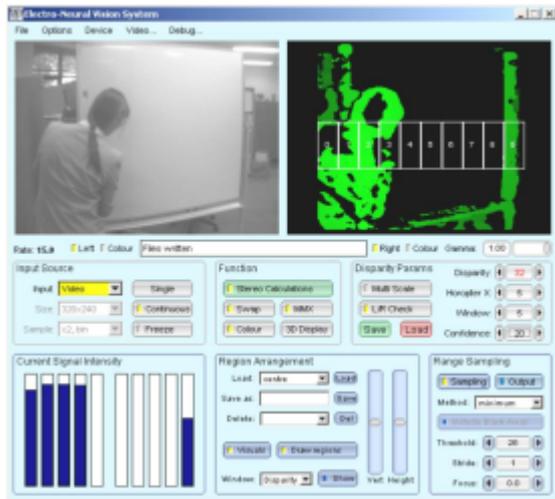


Figure 5. Disparity map of a featureless surface

3.2 Perceiving Landmarks with the ENVS

To enable landmarks to be perceived by blind users the ENVS is equipped with a GPS unit, a digital compass and a database of landmarks. The digital compass, shown in Figure 6, is mounted on the stereo camera headset and is used to determine if the user is looking in the direction of any landmarks.



Figure 6. Digital compass mounted on headset

Landmarks can be loaded from a file or entered by the user by pressing a button on the ENVS and are associated with their GPS location and an ID number. Landmarks are considered to be any significant object or feature in the environment that can enable the user to approximate his or her position from. They can also be a location the user wants to remember, e.g. a bus stop or a parking location of a vehicle. By using the GPS unit to obtain the user's location the ENVS is able to maintain a list of direction vectors to landmarks that are within a set radius from the user. The landmark radius can be set to short or long range (200-600m) by the user via a switch on the ENVS unit. When a landmark is calculated to be within the

user's ENVS visual field, (as determined by the headset compass and the set landmark radius), the perceived landmark's ID is encoded into a sequence of pulses and delivered to the user via the finger which represents the direction of the landmark. For example, if a landmark is determined to be in the far left visual field, the pulse sequence corresponding to the landmark will be felt on the left pinky finger.

To encode the landmark's ID a 5 bit sequence of dots and dashes carried by a 400Hz signal is used to represent binary numbers. To avoid interfering with the range readings of objects, which are also delivered to the fingers via the ENVS (see Section 3.1), locations are delivered to the fingers in 5 second intervals. Thus if a landmark is detected, the user will receive range readings via the fingers for 4 seconds followed by approximately 1 second of landmark ID information. If more than one landmark is present within the set landmark radius and the visual field of view, the landmark nearest to the center of the visual field will be output to the user.

By using 5 bits to represent landmark IDs, the ENVS is able to store up to 32 locations which proved adequate for our experiments. The distance to the landmark is indicated by the intensity of the pulses. Weak sensations indicate that the landmark is near to the set landmark radius. Strong sensations indicate that the landmark is only meters away from the user. If the user has difficulty recognizing landmarks by their pulse sequence a button is available on the ENVS unit to output the location, distance and bearing of the perceived landmark as speech.

4 Experimental Results

To test the ENVS we conducted a number of experiments within the campus grounds with different users to determine the extent to which the users could navigate the campus environment without any use of the eyes. All the ENVS users were familiar with the campus grounds and the landmarks stored in the ENVS and had no visual impairments. To simulate blindness with sighted users the stereo camera headset was designed to fit over the user's eyes so that no light whatsoever could enter the eyes. All reported tests were conducted on users who had at least 1 hour practise at using the ENVS.

4.1 Navigating the Car Park

Our first test was done to determine if the ENVS users could navigate a car park and arrive at a target vehicle location that was encoded into the ENVS as a landmark. Each user was fitted with the ENVS and lead blindfolded to a location in the car-park that was unknown to them and asked to navigate to the target vehicle using only the ENVS electro-tactile signals. The car-park was occupied to approximately 75% of its full capacity and also contained some obstacles comprising green strips and lighting poles.

Generally we found all users were able to perceive and describe their surroundings and the location of the target vehicle in sufficient detail for them to be able to navigate to the target vehicle without bumping into cars or lighting posts. We found experienced users could also interpret the range data without any need to hold their hands in front of the abdomen and could walk between closely spaced vehicles without colliding with the vehicles.

Figure 7a shows a user observing two closely spaced vehicles in the car-park with the ENVS. The profile of the space between the vehicles can be seen in the disparity map, shown in the top right of Figure 7b, and in the finger pulse histogram shown at the lower left of Fig 7b. The yellow bar at the left forefinger position of the finger histogram, indicates that target vehicle is located slightly to the left of where the user is looking and at a distance of approximately 120m.



(a)



(b)

Figure 7. (a) Photo of an ENVS user negotiating the car-park. (b) The ENVS screen dump showing the perceived vehicles and the target landmark's direction and distance.

4.2 Navigating the Campus

Experiments were also performed within the campus to determine the potential of the EVNS to enable

blindfolded users to navigate the campus without using other aids. The main test was to see if users could navigate between two locations some distance apart (~500m) and avoid any obstacles that might be in the way. All users who performed this test were familiar with the campus environment and had some experience at using the ENVS. The path was flat and contained no stairs between the two locations. A number of familiar landmarks were stored in the ENVS in the vicinity of the two locations.

We found that all users were able to avoid obstacles, report their approximate location and orientation and arrive at the destination without difficulty. Unlike the car park, the path was comprised of pavers with many joints making it clearly visible to the disparity cameras. Consequently, this delivered clearly defined range readings of the paved path to the user via the ENVS unit as shown in Figure 8.



(a)



(b)

Figure 8. (a) Photo of an ENVS user surveying a paved path in the campus environment. (b) The associated ENVS screen dump. The yellow bar in the range histogram indicates a landmark being delivered to the user.

This presented a minor problem for some inexperienced users as they were unable to determine where the path ended and the grass began in some

places using the ENVIS range stimulus alone. However, this did not cause any collisions and the users became quickly aware of the edge of the path whenever their feet made contact with the grass. We hope to overcome this minor problem by encoding colour into the range signals delivered to the fingers by varying the frequency of the tactile signals.

5 Conclusion

This paper describes our preliminary experimental results in developing a substitute vision system for the blind capable of providing 3D perception of the environment and landmark identification via an electro-tactile interface. The main problem with existing substitute vision systems for the blind is that the information delivered to the user is in a form that is either hard for the user to understand or difficult for the brain to derive a 3D model of the environment from. Also, existing substitute vision systems do not provide an adequate means for the user to perceive landmarks which are important for navigation. Although, commercial GPS units are available for the blind, they generally use an audio or voice interface which can interfere with a blind person's ability to hear important environmental sounds. The sounds made by talking GPS units can also undesirably attract the attention of nearby individuals.

Our outdoor experimental results demonstrate that our ENVIS is able to provide a blindfolded user with the ability to avoid obstacles, navigate the environment and locate his or her position in the environment via familiar landmarks by interpreting sensory data via the electro-tactile data gloves.

With further work we hope to develop the ENVIS into an effective device capable of providing the blind with increased environment perception and autonomy. This additional work includes the incorporation of colour perception into the range signals delivered to the user via the electro-tactile interface, adding infrared range sensors into the headset for detecting the range of featureless surfaces, the development of compact hardware for reducing the bulkiness of the ENVIS and the fabrication of alternative TENS interfaces eliminating the need for the user to wear electro-tactile data gloves.

6 Acknowledgements

This work was undertaken with the assistance of an Australian Research Council Discovery Grant.

7 References

[1] World Health Organisation, Magnitude and Causes of Visual Impairment, *Fact Sheet No.282*, November 2004.

[2] Wyatt, J.L. and Rizzo, J.F., Ocular Implants for the Blind, *IEEE Spectrum*, Vol.33, pp.47-53, May 1996.

[3] Rizzo, J.F. and Wyatt, J.L., Prospects for a Visual Prosthesis, *Neuroscientist*, Vol.3, pp.251-262, July 1997.

[4] Rizzo, J.F. and Wyatt, J.L., Retinal Prosthesis, in: *Age-Related Macular Degeneration*, J.W. Berger, S.L. Fine and M.G. Maguire, eds., Mosby, St. Louis, pp. 413 - 432. 1998.

[5] Dobbelle, W. Artificial Vision for the Blind by Connecting a Television Camera to the Visual Cortex, *American Society of Artificial Internal Organs Journal*, January/February 2000.

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

[7] Kay, L. Auditory Perception of Objects by Blind Persons Using Bioacoustic High Resolution Air Sonar. *JASA*, Vol 107, pp 3266-3275, No 6, June 2000.

[8] Kaczmarek, K.A. and Bach-y-Rita, P., *Tactile Displays, in Virtual Environments and Advanced Interface Design*, Barfield, W. and Furness, T., Eds. New York: Oxford University Press, pp. 349-414, 1995.

[9] Helal, A.; Moore, S.E.; Ramachandran, B., Drishti: an integrated navigation system for visually impaired and disabled, *Proceedings, Fifth International Symposium on Wearable Computers*, 2001.

[10] Mosen, J.; What's New with Pulse Data International's Brailnote PDA for the Blind, *Center on Disabilities Technology and Persons with Disabilities Conference*, 2005.

[11] Daviault, J.; Study On The Benefits Of GPS-Based Orientation Tools for the Mobility and Orientation of the Blind, *Center on Disabilities Technology and Persons with Disabilities Conference*, 2005.

[12] Balachandran, W.; Cecelja, F.; Ptasinski, P., A GPS based navigation aid for the blind, *ICECom*, 2003

[13] Virkes, D.; Nenadic, K.; Virkes, D.Z., Positioning systems for mobility of the blind, overview and analysis, *Proceedings, 46th International Symposium Electronics in Marine (ELMAR)*, 2004.

[14] S. Meers and K. Ward, A Vision System for Providing 3D Perception of the Environment via Transcutaneous Electro-Neural Stimulation. *Proceedings of the 8th IEEE International Conference on Information Visualisation*, London, pages 546-552, July 2004.

[15] Banks, J. Bennamoun, M. and Corke, P., Non-Parametric Techniques for Fast and Robust Stereo Matching, *Proceedings of the IEEE TENCON'97*, Brisbane, December 1997.