

# Haptic Gaze-Tracking Based Perception of Graphical User Interfaces

Simon Meers, Koren Ward  
University of Wollongong  
meers@uow.edu.au, koren@uow.edu.au

## Abstract

*This paper presents a novel human-computer interface that enables the computer display to be perceived without any use of the eyes. Our system works by tracking the user's head position and orientation to obtain their 'gaze' point on a virtual screen, and by indicating to the user what object is present at the gaze location via haptic feedback to the fingers and synthetic speech or Braille text. This is achieved by using the haptic vibration frequency delivered to the fingers to indicate the type of screen object at the gaze position, and the vibration amplitude to indicate the screen object's window-layer, when the object is contained in overlapping windows. Also, objects that are gazed at momentarily have their name output to the user via a Braille display or synthetic speech. Our experiments have shown that by browsing over the screen and receiving haptic and voice (or Braille) feedback in this manner, the user is able to acquire a mental two-dimensional representation of the virtual screen and its content without any use of the eyes. This form of blind screen perception can then be used to locate screen objects and controls and manipulate them with the mouse or via gaze control. Our experimental results are provided in which we demonstrate how this form of blind screen perception can effectively be used to exercise point-and-click and drag-and-drop control of desktop objects and open windows by using the mouse, or the user's head pose, without any use of the eyes.*

## 1 Introduction

To access the computer, most blind users use a screen reader program, like JAWS [4] or Window-Eyes [6], equipped with synthetic speech and/or a Braille display [3, 21, 25]. This gives blind users access to the text appearing on the screen. However, screen readers do not provide blind users with the ability to perceive and access Graphical User Interface (GUI) controls by using the mouse, which requires considerable hand-eye coordination to be used effectively. Consequently, most control access to the computer by the blind is achieved by using command line interfaces and/or by using specific combinations of control keys. This can limit the applications available to blind users as well as the tasks that can be performed on cer-

tain GUI-based applications. For example, when the software development industry changed from command line compilers to integrated development environments, almost all blind programmers were made redundant due their inability to effectively perceive and use such GUI development environments and the difficulty adapting such environments to their needs [1, 22].

Tactile devices for enabling blind users to perceive graphics or images on the computer have been under development for some time. For example the haptic mouse (e.g. [7, 9, 24]) can produce characteristic vibrations when the mouse cursor is moved over screen icons, window controls and application windows. Although this can indicate to a blind user what is currently beneath the mouse cursor, it does not give a blind user much idea of where screen objects are located on the screen or where the mouse is currently located.

Force feedback devices, like the PHANToM [5], and tactile (or electro-tactile) displays, eg. [8, 10, 11, 13], can enable three-dimensional graphical models or two-dimensional black-and-white images to be visualised by using the sense of touch (see [2] for a survey). However, little success has been demonstrated with these devices toward enabling blind users to interact with typical GUIs other than some simple memory and visualisation experiments like "the memory house" [23] which involves discovering buttons on different planes via force feedback and remembering the buttons that play the same sounds when found.

Refreshable haptic Braille displays comprised of a two-dimensional matrix of raisable pins have been used to devise some GUI controls. For example HyperBraille [12] has adopted the concept of pull-down menus which are customised to suit Braille displays. Also, the ACCESS project [20] attempts to present web based hypermedia information in Braille format within a refreshable Braille display. However, these systems do not provide the user with any two-dimensional perception of the screen content which can result in too little graphical information being presented to the user for effective GUI interaction. Also, confusion can occur where context-sensitive menus are involved. Furthermore, the main GUI concepts of point-

and-click and drag-and-drop do not implement well within Braille displays due to their limited size and the hands being occupied reading through Braille for much of the time.

In an attempt to address these deficiencies, we have been developing a gaze-tracking interface for enabling blind users to perform many of the tasks that sighted users can perform on typical GUIs. Our system works by tracking the user's head pose with gaze-tracking hardware [14] to obtain the head's gaze position on a large virtual screen and by indicating to the user what is present at the gaze location via vibro-tactile or electro-tactile feedback to the fingers. The haptic vibration frequency felt by the fingers indicates the type of screen object at the gaze position. Likewise, the amplitude of the haptic vibrations indicates the window-layer of the screen object when the object is contained in overlapping windows. The mouse position is revealed with high amplitude vibrations when it is gazed at by the user. Any screen object's name can also be output to the user via synthetic speech or a Braille display when the user gazes at the object momentarily.

By browsing over the screen and receiving haptic and voice (or Braille) feedback in this manner, the user is able to quickly acquire a mental representation of the virtual screen and its content. We have found this to be sufficient for performing many point-and-click or drag-and-drop tasks which are inherent in all GUI applications and operating systems. In the following sections we firstly provide a brief background review of our blind vision work leading up to the development of our haptic gaze-tracking user interface. This is followed with the implementation details of our system and our experimental results.

## 2 Background

In previous work [15, 16], we have been investigating the use of electro-tactile user interfaces and range sensing devices for providing depth and colour perception of the environment to the blind. Our vision system, shown in Figure 1, works by extracting depth and colour information from sensors mounted on a headset worn by the user, and delivering this information to the fingers via electro-tactile stimulation. To interpret the range and colour data, the user simply imagines that their hands are held with fingers extended in the direction viewed by the cameras. Each finger receives feedback regarding the area of the environment it is envisaged to be pointing toward. The intensity of electro-tactile stimulation felt indicates the distance to objects, while the frequency indicates the predominate colour of each region.

Our experiments have shown that this continuous delivery of environmental depth and colour information to the user, in a form that is easy to interpret, enables the user to realise the 3D profile of their surroundings as well as the presence of landmarks based on their shape and colour, by



Figure 1: The Electro-Neural Vision System Prototype

surveying the environment with the sensor headset. Our experimental results demonstrate that this form of perception makes it possible for the user to navigate certain environments, recognise their location, and perceive the size, colour and movement of objects within the surrounding environment without any visual perception from the eyes.

To adapt this concept to interpreting screen content on the computer we use a similar method for providing electro-tactile feedback to the fingers. However, instead of using a sensor headset, we use a head pose tracking system [14] that we developed to determine the user's gaze position on a predefined *virtual screen*. We also developed software for interpreting the virtual screen's content at the gaze position and delivering this information to the fingers via electro-tactile stimulation. The prototype electro-tactile interface is shown in Figure 2. This is comprised of Transcutaneous Electro-Neural Stimulation (TENS) electrodes that are fitted to the fingers and connected to a purpose-built computer-controlled TENS unit (not shown). In the following section, we provide further details of our haptic gaze-tracking user interface and the design of a vibro-tactile keyboard that could potentially be used instead of the electro-tactile interface.

## 3 Gaze-Tracking Interface

The haptic gaze-tracking user interface is basically a device for enabling the computer screen to be perceived without any use of the eyes. This is achieved by tracking the head pose of the user and by delivering haptic feedback to the user's fingers which indicates what is located at the gaze position on a predefined virtual screen. This effectively enables the user to browse over the virtual screen and mentally visualise its content. It also makes it possible to perceive the mouse position on the virtual screen and to



Figure 2: Electro-Tactile Interface comprised of Transcutaneous Electrodes fitted to the hands.

manipulate icons, menus and other GUI controls through normal use of the mouse. The user can also choose to move the mouse with their gaze position (rather than independently) if desired which can simplify GUI-hand interaction considerably. Figure 3 shows the main hardware components of our haptic gaze-tracking user interface. Haptic feedback is provided by the TENS electrodes fitted to the user's fingers and connected the TENS unit (not shown). The head pose tracking system is comprised of the USB camera and the spectacles fitted to the user. The spectacles contain three infrared LEDs that are tracked by image processing software we developed. Note: the computer monitor shown in Figure 3, supporting the USB camera, is unnecessary for the system and was left turned off to conduct our experiments.



Figure 3: Hardware Components: Gaze-Tracking Spectacles/Camera and Electro-Tactile Interface

### 3.1 The Virtual Screen

For the system to work effectively, considerable attention had to be paid to appropriately defining the perceived virtual screen, its content and how this information is encoded for haptic interpretation. To provide increased resolution to the user we defined a large virtual screen that was divided up into a matrix of cells. A visual representation of part of the virtual screen is depicted in Figure 4.

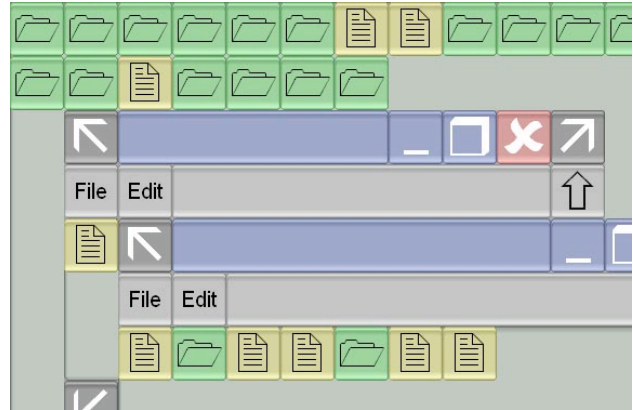


Figure 4: Close up representation of the virtual screen.

To represent the desktop environment, each cell on the virtual screen is considered a unit of perception and may contain an icon or a GUI control. The size of the virtual screen is defined by the gaze area about which the user can comfortably browse. The cell size, which defines the virtual screen resolution, can be set by the user and would typically be smaller for experienced users (higher resolution) than for novice users (lower resolution). For our experiments we defined the virtual screen to be 6m wide x 4m high and at a distance of 2m from user. The size of each cell was set to 250mm x 250mm which provides a resolution of 24 cells wide x 16 cells high. Figure 5 shows a photo of our system together with an artist's representation of the virtual screen being perceived by the user.

Our haptic gaze-tracking interface is intended to be mainly implemented at the operating system level as a means of providing alternative access to the desktop and typical GUI-based applications by the blind. Our prototype and initial experiments were based upon the Microsoft Windows desktop environment due to it being the most extensively deployed operating system. Subsequently, the main task involved reconstructing all of the GUI components of the Windows desktop (e.g. taskbar, icons, windows and controls) to cells comprising our virtual screen. Figures 4 and 5 show an example of how this was done with two open windows and some desktop icons.

Our use of overlapping, resizable windows with scrollbars, etc., is not indicative of a belief that this is necessarily the best interface for blind users, however, this form

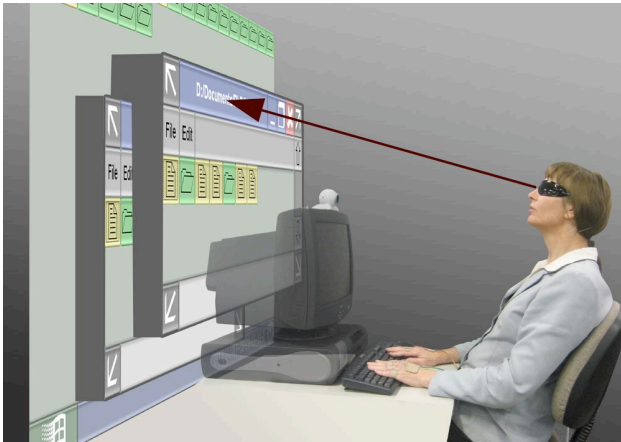


Figure 5: The Haptic Gaze-Tracking User Interface and an artist's impression of the Virtual Screen

of interface has been widely adopted by operating system providers and application developers, and can be difficult for the blind to use due to the extensive use of GUI controls. Such GUI software applications can also be difficult to adapt to the needs of the blind by using development tools such as Microsoft's Active Accessibility [17]. Our haptic gaze-tracking technique may allow GUIs to be made accessible to blind users without significant redevelopment work. It may also provide the blind with an effective means of perceiving screen content and operating GUI controls via point-and-click and drag-and-drop actions which form the basis of nearly all modern computer applications.

To implement the virtual screen, icons on the desktop are mapped to corresponding cells on the virtual screen and are made to snap to the nearest cell when released from the mouse. The taskbar is similarly mapped to the bottom row of cells on the virtual screen. Desktop icons are assigned a haptic type that defines the vibrations produced when the icon is gazed at. For our preliminary experiments, we simply defined five haptic types which were represented with the frequencies 10, 20, 40, 80 and 160 Hz. These haptic types correspond to "file", "folder", "application", "control" and "empty cell" respectively. (Note: with the use of modulated frequencies this could be expanded.) Cells are also attributed to one of four levels of haptic intensity to enable the user to determine if the cell is at the desktop layer (lowest intensity); within an inactive window in front of the desktop but below another window (low-medium intensity); within the frontmost window (high-medium intensity); or occupied by the mouse (highest intensity). This can be seen in the visual representations of the virtual screen shown in Figures 4 and 5.

Our experiments have shown that by browsing over the virtual screen and receiving haptic feedback in this manner the user is able to perceive the screen content in terms

of where the mouse is currently located and where various icon types and controls are located on the desktop and within open windows. By gazing at any icon, control or label momentarily, the object's name is spoken via synthetic speech. (Note: Alternatively, the object's name can be displayed on a Braille display if one is fitted to the computer.) Consequently, it is relatively straightforward to perform tasks such as: locating a specific file or folder; opening files, folders or applications; or moving objects to another cell or folder by using the mouse or via gaze control. When a window is opened to reveal an open folder (or application window), its edges can be detected by the changes in haptic intensity when the user browses over it. This also enables the window controls to be located with relative ease for resizing, closing or minimising the window by using the mouse under gaze or hand control.

Our haptic gaze-tracking interface is also equipped with an option for increasing the user's perceived area of the virtual screen near the gaze position. This is achieved by using vibrations delivered to the index fingers to reveal the cell at the gaze position and the vibrations delivered to the three outer fingers of each hand (ie. the pinky, ring and middle fingers) to reveal the content of three cell positions to the left and right of the gaze position as depicted in Figure 6.

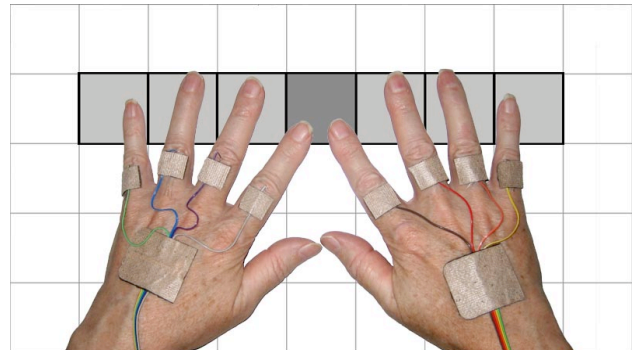


Figure 6: The electro-tactile interface showing how each finger is mapped to the gazed area of the virtual screen

The overall intensity delivered to each finger can be pre-set in the system's setup. Normally, more intensity is desirable at the index fingers (cf. foveal vision) so that changes occurring at the outer (peripheral) fingers do not interfere with perceiving the object at the gaze centre. This enables the user to scan the entire virtual screen more quickly to reveal the location of screen icons, controls and the mouse. Our experiments have shown that even though it can be difficult to simultaneously realise the haptic types of all the screen objects that are browsed over rapidly using this form of foveal perception, it enables the user to quickly form a mental occupancy map of the screen layout. The user can then gaze at the objects of interest directly to exam-

ine them in more detail with the central (foveal) perception and retrieve additional information via speech synthesis or Braille output if desired.

### 3.2 Interacting with the Grid

To assist inexperienced users in navigating the grid-based interface, an optional auditory ‘beep’ cue is provided when the gaze moves from one cell to another. This is particularly useful when the user is scanning adjacent cells containing elements of the same type, since the haptic output will not change in this situation, making it difficult to perceive where one cell ends and the next begins.

Our initial experiments with the interface also revealed confusion could occur when the user’s gaze travelled along cell borders. For example, Figure 7(a) shows a path that the user’s gaze could take across the virtual screen and the order in which the cells are visited. If these cells contain different screen objects the corresponding haptic responses would be similarly varying and confusing. To address this problem we introduced a threshold algorithm which prevented the gaze point from visiting a new cell until the gaze point moved more than one-quarter of a cell’s width beyond the cell border (see Figure 7(b)). This measure enabled neighbouring cells to be differentiated and resolved more easily improving the interface’s usability.

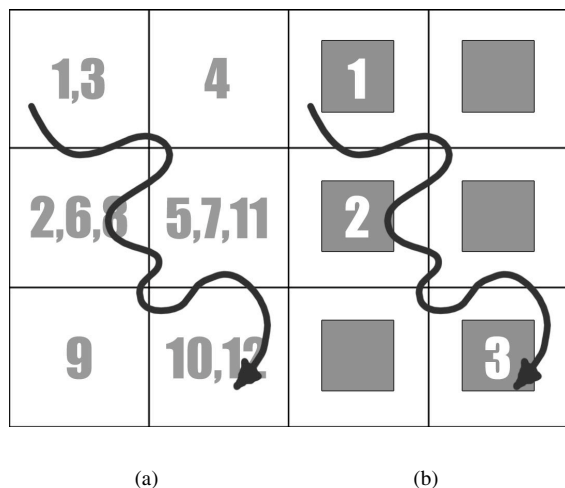


Figure 7: Grid/gaze interaction methods

### 3.3 Vibro-Tactile Interface

At this stage our experiments have been conducted with an electro-tactile interface for providing haptic feedback to the user’s fingers. Although this has proved effective it requires electrodes to be fitted to the user’s fingers and wires from the electrodes to be connected to the TENS unit. To overcome this requirement we are developing wireless TENS electrode assemblies that can be fitted to the hands

(or elsewhere) with relative ease. We are also considering the use of vibro-tactile feedback as an alternative means of providing haptic feedback to the user. Figure 8 shows an artist’s impression of one possible means of incorporating vibro-tactile feedback into a conventional computer keyboard. Here vibro-tactile actuators (or tactors) are used to provide feedback to the outer fingers, for providing peripheral perception, and a small Braille display is provided to inform the user of what is located at the gaze (or foveal) position via the pointer fingers. The mouse is implemented as a trackball that is controlled by the thumbs. The user can also translate (or morph) the mouse cursor to the gaze position by pressing a third mouse button. We expect to have our vibro-tactile interface and associated experimental results completed and published at a later stage.

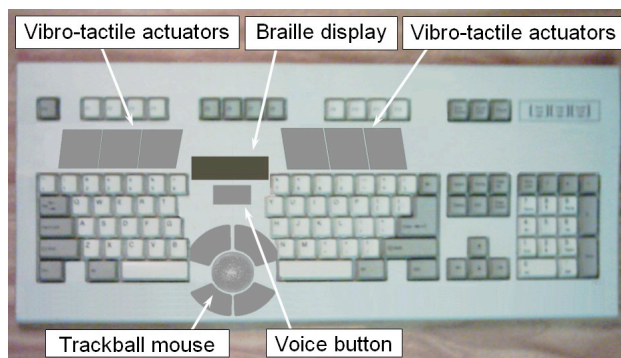


Figure 8: Artist’s impression of the haptic keyboard.

### 3.4 Gaze-Tracking System

Although a variety of gaze-tracking systems are commercially available, we found them all to be too computationally expensive or inaccurate for our application. Consequently, for this application we implemented an inexpensive and robust method for tracking the head position and orientation of the user by using a single low-cost USB camera and infrared light-emitting diodes concealed within spectacle frames worn by the user [14]. Unlike gaze and head pose tracking systems which rely on high-resolution stereo cameras and complex image processing hardware and software to find and track facial features on the user (eg. [18, 19, 26]), our head pose tracking system is able to efficiently locate and track the head’s orientation and distance relative to the camera with little processing. Due to the infrared light-emitting diodes having fixed geometry, the system does not have to contend with the varying facial features of different users and therefore does not require any calibration procedure or training to accommodate any user. Furthermore, this system is unaffected by varying lighting conditions and can be used in the dark. The spectacles can be replaced by any other head-mounted object if desired. Our tests have shown that our head pose track-

ing system is accurate to within 0.5 degrees when the user is within one metre of the camera. This compares well with more expensive head tracking systems and proved adequate for our haptic gaze-tracking interface.

### 3.5 Learning Procedure

We found that a setup procedure and practice was necessary for new users to be able to use the haptic gaze-tracking user interface effectively. The setup procedure was necessary to ensure that the electro-tactile signals delivered to the fingers were considered comfortable and appropriate for the user. Once set up, these settings are able to be saved to a file and restored whenever the user logs in. Also, practice at identifying desktop icons and window controls with the peripheral output turned off was necessary in order to avoid any confusion from trying to interpret too much information simultaneously with the hands. Once the user was able to associate the different electro-tactile signals with the different icon types, controls and window layers, the peripheral fingers could be gradually turned on and utilised for perceiving screen content more quickly.

## 4 Experimental Results

Our experiments with the haptic gaze-tracking user interface have demonstrated that by browsing over the virtual screen and receiving haptic feedback in this manner the user is able to perceive and remember the types and locations of screen icons, the mouse and any GUI controls that are located on the desktop (or within open windows) without any use of the eyes. Furthermore, by gazing at any icon, control or label momentarily, and by having the object's name spoken via synthetic speech, it was relatively easy for the user to perform tasks such as: finding and starting applications, locating a specific file or folder, or opening files or folders. The user could also perform drag-and-drop manipulation of such objects, such as moving screen icons onto another cell or into another window by using the mouse with the hand or under gaze control. When a window was opened to reveal an open folder or application window, its edges could be easily detected by the changes in haptic intensity when the user browsed over it. This also enabled the window controls and menus to be located with relative ease and manipulation tasks to be performed like resizing, closing or minimising windows by using the mouse under gaze or hand control.

### Conclusion

With GUI-based computer applications used extensively in the workplace, for education and for leisure, the blind are becoming increasingly disadvantaged and handicapped at performing many tasks on the computer that sighted users take for granted. Although efforts have been made to provide increased computer accessibility to the blind (e.g. Microsoft Active Accessibility [17] and JAWS

for Windows [4]), many applications and subsequent occupations remain inaccessible and unachievable for the blind. This paper presents a novel haptic gaze-tracking human-computer interface that enables the Windows desktop environment to be graphically perceived and accessed without any use of the eyes. Our experimental results have demonstrated that significant blind GUI perception and interactivity is able to be performed with our system including point-and-click and drag-and-drop control of conventional GUI objects. To further develop this work we intend also applying this method for achieving blind GUI interactivity to typical GUI-based computer applications such as web-browsers and word processors.

### References

- [1] Alexander, S. Blind Programmers Facing Windows. *Computer World*, November 2, 1998. Reprinted online by CNN: <http://www.cnn.com/TECH/computing/9811/06/blindprog.idg/>
- [2] Chouvardas, V., Miliou, A. and Hatalis, M., Tactile Displays: a short overview and recent Developments, *5th International Conference on Technology and Automation*, (October, Thessaloniki, Greece), Pages 246-251, 2005.
- [3] Freedom Scientific. Focus 40-80 Braille Display. [http://www.freedomscientific.com/fs\\_products/displays\\_focus40-80.asp](http://www.freedomscientific.com/fs_products/displays_focus40-80.asp)
- [4] Freedom Scientific. Job Access With Speech (JAWS). [http://www.freedomscientific.com/fs\\_products/software\\_jaws.asp](http://www.freedomscientific.com/fs_products/software_jaws.asp)
- [5] Fritz, J.P., and Barner, K. E. Design of a Haptic Visualization System for People with Visual Impairments. *IEEE Transactions on Rehabilitation Engineering*, vol. 7, No 3, 1999, pp 372-384.
- [6] GW Micro. Window-Eyes. <http://www.gwmicro.com/>
- [7] Hughes, R.G. and Forrest, A.R. Perceptualisation using a tactile mouse. In *Proceedings of the 7th conference on Visualization '96*. Los Alamitos, CA, USA: IEEE Computer Society Press, 1996, pp. 181-ff.
- [8] Ikei, Y., Wakamatsu, K., and Fukuda, S. Texture display for tactile sensation. In *Proceedings of the Seventh International Conference on Human-Computer Interaction*, ser. Virtual Reality, vol. 2, 1997, pp. 961-964.
- [9] Immersion Corporation. iFeel Mouse. <http://www.immersion.com>

- [10] Kaczmarek, K., Webster, J., Bach-y-Rita, P. and Tompkins, W. Electrotactile and vibrotactile displays for sensory substitution systems. *IEEE Transactions on Biomedical Engineering*, vol. 38, Issue 1, pp. 1 - 16, 1991.
- [11] Kawai, Y. and Tomita, F. Interactive tactile display system: a support system for the visually disabled to recognize 3d objects. In *Proceedings of the second annual ACM conference on Assistive technologies*. ACM Press, 1996, pp. 45-50.
- [12] Kieninger, T. The “growing up” of hyperbraille an office workspace for blind people. In *Proceedings of the 9th annual ACM symposium on User interface software and technology*. New York, NY, USA: ACM Press, 1996, pp. 67-73.
- [13] Maucher, T., Meier, K., and Schemmel, J. The heidelberg tactile vision substitution system. In *Proceeding of the Sixth International Conference on Tactile Aids, Hearing Aids and Cochlear Implants*, 2000.
- [14] Meers, S., Ward, K. and Piper, I. Simple, Robust and Accurate head pose Tracking Using a Single Camera. Accepted in *Proceedings of the Thirteenth Annual Conference on Mechatronics and Machine Vision in Practice*, (December, Toowoomba, Australia), 2006.
- [15] Meers, S. and Ward, K. A Vision System for Providing 3D Perception of the Environment via Transcutaneous Electro-Neural Stimulation. In *Proceedings of the 8th IEEE International Conference on Information Visualisation* (July, London, UK), 2004, pp.546-552.
- [16] Meers, S. and Ward, K. A Vision System for Providing the Blind with 3D Colour Perception of the Environment. In *Proceedings of the 2005 Asia-Pacific Workshop on Visual Information Processing* (November, Hong Hong), 2005, pp.102-108.
- [17] Microsoft Corporation. Microsoft Active Accessibility. <http://msdn.microsoft.com/library/default.asp?url=/library/en-us/dnanchor/html/accessibility.asp>
- [18] NaturalPoint Inc. TrackIR, 2006. <http://www.naturalpoint.com/trackir>
- [19] Newman, R., Matsumoto, Y., Rougeaux, S., and Zelinsky, A. Real-time stereo tracking for head pose and gaze estimation. In *Proceedings. Fourth IEEE International Conference on Automatic Face and Gesture Recognition*, pages 122-128, 2000.
- [20] Petrie, H., Morley, S., McNally, P., Graziana, P. and Emiliani, P. Access to hyper-media systems for blind people. Sensort Disabilities Research Unit, 1996.
- [21] Roberts, J., Slattery, O., and Kardos, D. Rotating-wheel braille display for continuous refreshable braille. In *Society for Information Display conference in Long Beach, California, May 18; 2000 SID International Symposium Digest of Technical Papers*, vol. XXXI, May 2000, pp. 1130-1133.
- [22] Siegfried, R., Teaching the Blind to Program Visually, In *Proceedings of ISECON 2004*, Newport, December 2004.
- [23] Sjöström, C. Designing Haptic Computer Interfaces For Blind People. In *Proceedings of the Sixth IEEE International Symposium on Signal Processing and its Applications*, Kuala Lumpur, Malaysia, August 13 - 16, 2001.
- [24] Virtouch Ltd. Virtouch Imaging Applications. Jerusalem, Israel. <http://www.virtouch2.com>
- [25] Yobas, Y., Durand, D.M., Skebe, G.G., Lisy, F.J. and Huff, M.A. A novel integrable microvalve for refreshable braille display system. In *Journal Of Microelectromechanical Systems*, June 2003, vol. 12, no. 3, pp. 252 - 263.
- [26] Zhu, Z. and Ji, Q.. Real time 3d face pose tracking from an uncalibrated camera. In *First IEEE Workshop on Face Processing in Video, in conjunction with IEEE International Conference on Computer Vision and Pattern Recognition (CVPR'04)*, (June, Washington DC), 2004.