Head-Tracking Haptic Computer Interface for the Blind

Simon Meers and Koren Ward
University of Wollongong
Australia

1. Introduction

In today's heavily technology-dependent society, blind and visually impaired people are becoming increasingly disadvantaged in terms of access to media, information, electronic commerce, communications and social networks. Not only are computers becoming more widely used in general, but their dependence on visual output is increasing, extending the technology further out of reach for those without sight. For example, blindness was less of an obstacle for programmers when command-line interfaces were more commonplace, but with the introduction of Graphical User Interfaces (GUIs) for both development and final applications, many blind programmers were made redundant (Alexander, 1998; Siegfried et al., 2004). Not only are images, video and animation heavily entrenched in today's interfaces, but the visual layout of the interfaces themselves hold important information which is inaccessible to sightless users with existing accessibility technology.

Screen reader applications, such as JAWS (Freedom Scientific, 2009b) and Window-Eyes (GW Micro, 2009), are widely used by the visually impaired for reading screen content and interpreting GUIs (Freitas & Kouroupetroglou, 2008). Although they allow the user to access the computer via control key commands and by listening to synthetic speech they can be slow and difficult for inexperienced users to learn. For example, JAWS has over 400 control key sequences to remember for controlling of the screen reader alone (Freedom Scientific, 2009a). Furthermore, a large amount of layout and format information is lost in the conversion from what is effectively a two-dimensional graphical interface into a linear sequence of spoken words.

Various interfaces have been devised which utilise force-feedback devices such as the PHANTOM (Massie & Salisbury, 1994), or (electro-)tactile displays (e.g. Ikei et al., 1997; Kaczmarek et al., 1991; Kawai & Tomita, 1996; Maucher et al., 2001) for haptic perception of three-dimensional models or simple two-dimensional images. Researchers such as Sjöström (2001) have demonstrated success with enabling blind users to interact with certain custom-built interfaces, but not typical GUIs.

Vibro-tactile devices such as the tactile mouse (Immersion Corporation, 2009; Hughes & Forrest, 1996; Gouzman & Karasin, 2004) are designed to provide characteristic tactile feedback based on the type of element at the mouse pointer location. Although a tactile mouse can give a blind user some sense of the spatial layout of screen elements, the inability of blind users to perceive exactly where the mouse pointer is located makes this form of interface ineffective for locating and manipulating screen elements.

Refreshable Braille displays have significantly higher communication resolution, and present information in a manner which is more intuitive for blind users, including the ability to represent text directly. Several projects have been undertaken to represent graphical interfaces using such displays. For example, HyperBraille (Kieninger, 1996) maps HyperText Markup Language (HTML) pages into Braille "pull down menu" interfaces. Recently, Rotard et al. (2008) have developed a web browser extension which utilises a larger pin-based tactile display with the ability to render simple images using edge-detection, as well as Braille representations of textual content. Such systems provide advantages beyond simple screen readers, but are still very limited in terms of speed of perception, layout retention and navigability.

To address these shortcomings we have been devising various interfaces for the visually impaired which involve head-pose tracking and haptic feedback. Our system utilises a head-pose tracking system for manipulating the mouse pointer with the user's 'gaze' which allows the user's hands to be free for typing and tactile perception.

This is implemented by mapping the graphical interface onto a large 'virtual screen' and projecting the 'gaze' point of the user onto the virtual screen. The element(s) at the 'focal' position are interpreted via tactile or voice feedback to the user. Consequently, by gazing over the virtual screen, the user can quickly acquire a mental map of the screen's layout and the location of screen elements (see Figure 1). By gazing momentarily at a single element, additional details can be communicated using synthetic speech output via the speakers or Braille text via a Braille display.

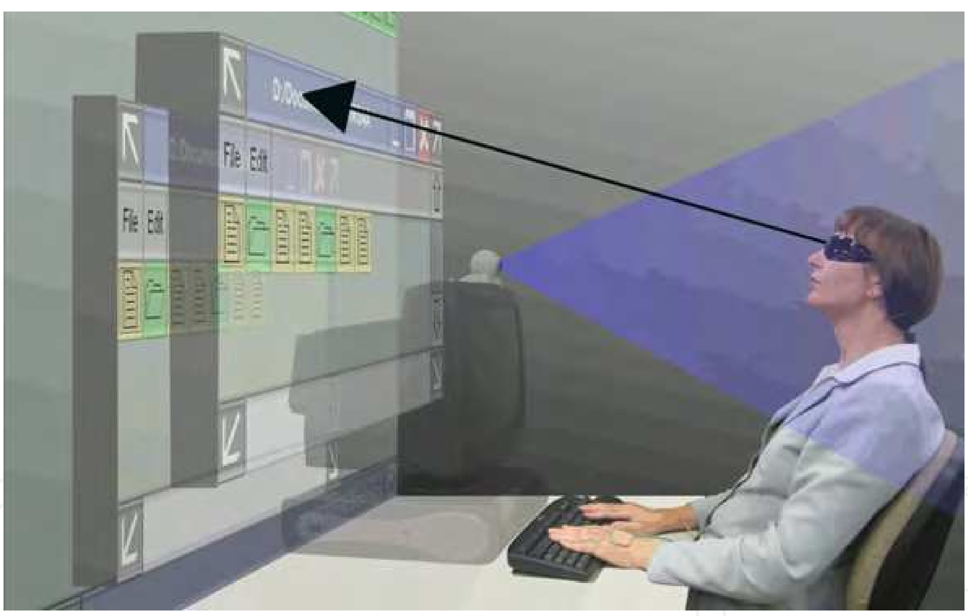

Fig. 1. Visual representation of the gaze-tracking "virtual screen" concept

We have experimented with a number of methods of mapping various graphical interfaces to blind gaze tracking virtual screens as well as a number of different haptic feedback devices. Details of these mapping techniques and haptic feedback devices are provided in the following sections.

2. Background

This project stems from our development of the Electro-Neural Vision System (ENVS) (Meers & Ward, 2004) – a device which allows its wearer to perceive the three-dimensional profile of their surrounding environment via Transcutaneous Electro-Neural Stimulation (TENS) and therefore navigate without sight or other aids. It utilises a head-mounted range-sensing device such as stereo cameras or an array of infrared sensors, pointed in the direction of the wearer's 'gaze'. The acquired depth map is divided into sample regions, each of which is mapped to a corresponding finger which receives electro-tactile stimulation of intensity proportional to the distance measured in that region. The frequency of the signal was used for encoding additional information such as colour (Meers & Ward, 2005b) or GPS landmark information (Meers & Ward, 2005a). Our experiments showed that this form of perception made it possible for unsighted users to navigate known environments by perceiving objects and identifying landmarks based on their size and colour. Similarly, unknown outdoor environments could be navigated by perceiving landmarks via GPS rather than their size and colour. Figure 2 shows the ENVS in use.

Fig. 2. Electro-Neural Vision System Prototype

Our ENVS experiments inspired us to implement a similar form of perception for interpreting the content of the computer screen. In this case, a large *virtual screen* was located in front of the user and a head-pose tracking system was used to track the 'gaze' position of the user on the virtual screen. To determine what is located at the gaze position on the virtual screen, pre-coded haptic feedback signals are delivered to the fingers via electro-tactile electrodes, a haptic keyboard or a refreshable Braille display. The following sections provide details of the head-pose tracking systems and haptic feedback devices deployed on our interface.

3. Gaze-Tracking Haptic Interface

The primary goal of our gaze-tracking haptic interface is to maintain the spatial layout of the interface so that the user can perceive and interact with it in two-dimensions as it was intended, rather than enforcing linearisation, with the loss of spatial and format data, as is the case with screen readers. In order to maintain spatial awareness, the user must be able to control the "region of interest" and understand its location within the interface as a whole. Given that we wanted to keep the hands free for typing and perception, the use of the head as a pointing device was an obvious choice – a natural and intuitive pan/tilt input device which is easy to control and track for the user (unlike mouse devices).

3.1 Head-pose tracking

While there are quite a number of head-pose tracking systems commercially available, we found that they were all either too cumbersome, computationally expensive or inaccurate for our requirements. Consequently, we developed our initial prototype using our own custom-developed head-pose tracker (Meers et al., 2006) which utilised a simple USB web camera and a pair of spectacles with three infrared LEDs to simplify the tracking process. This proved to be robust and accurate to within 0.5°.

To avoid the need for the user to wear special gaze-tracking spectacles, we developed a head-pose tracking system based on a time-of-flight camera (Meers & Ward, 2008). This not only made our interface less cumbersome to set up, but also provided the advantage of in-built face recognition (Meers & Ward, 2009) for loading user preferences, etc.

3.2 The Virtual Screen

Once the user's head-pose is determined, a vector is projected through space to determine the gaze position on the virtual screen. The main problem is in deciding what comprises a screen element, how screen elements can be interpreted quickly and the manner by which the user's gaze passes from one screen element to another. We have tested two approaches to solving these problems as explained in the following sections

3.2.1 Gridded Desktop Interface

Our initial experiments involved the simulation of a typical "desktop" interface, comprising a grid of file/directory/application icons at the "desktop" level, with cascading resizable windows able to "float" over the desktop (see Figure 3). The level of the window being perceived (from frontmost window to desktop-level) was mapped to the intensity of haptic feedback provided to the corresponding finger, so that "depth" could be conveyed in a similar fashion to the ENVS. The frequency of haptic feedback was used to convey the *type* of element being perceived (file/folder/application/control/empty cell). Figure 4 illustrates the mapping between adjacent grid cells and the user's fingers. The index fingers were used to perceive the element at the gaze point, while adjacent fingers were optionally mapped to neighbouring elements to provide a form of peripheral perception. This was found to enable the user to quickly acquire a mental map of the desktop layout and content. By gazing momentarily at an individual element, the user could acquire additional details such as the file name, control type, etc. via synthetic speech output or Braille text on a Braille display.

A problem discovered early in experimentation with this interface was the confusion caused when the user's gaze meandered back and forth across cell boundaries, as shown in Figure 5. To overcome this problem, a subtle auditory cue was provided when the gaze crossed boundaries to make the user aware of the grid positioning, which also helped to distinguish con-

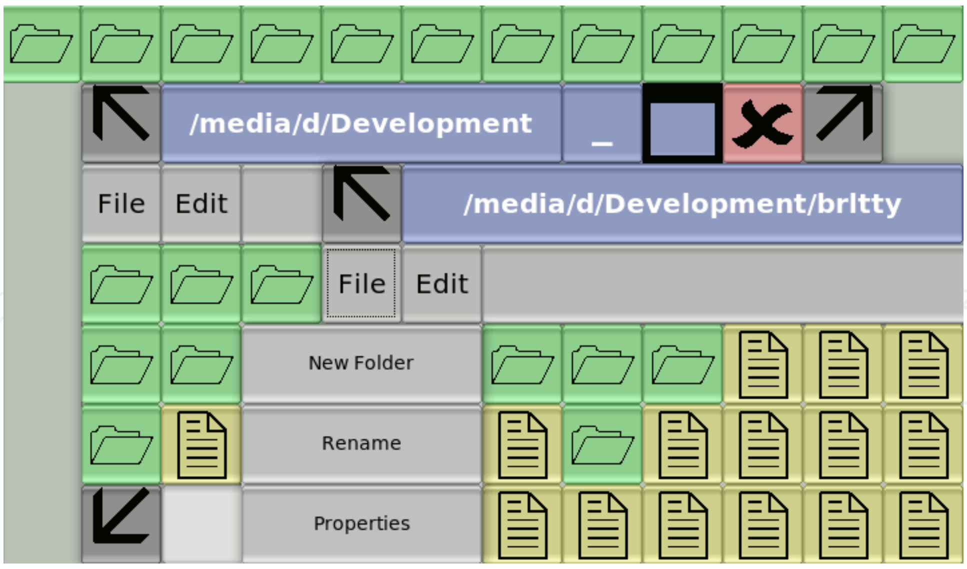

Fig. 3. Experimental desktop grid interface

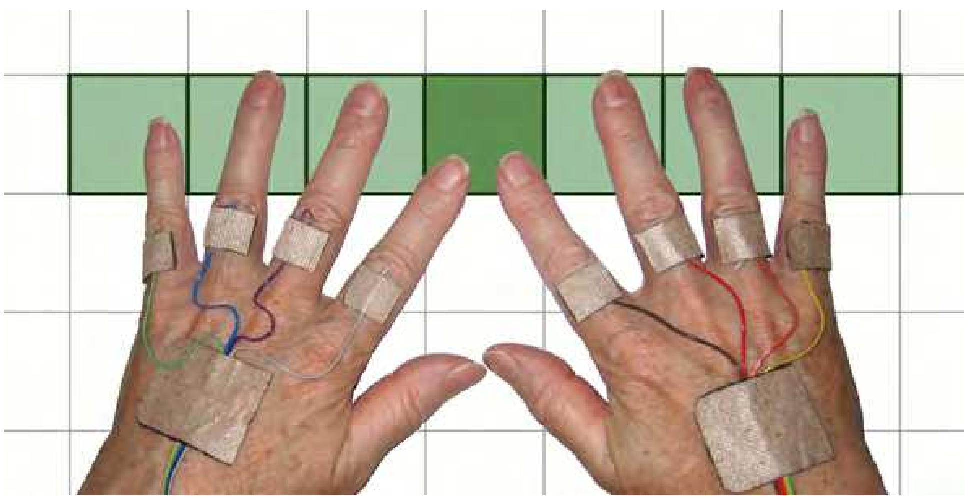
Fig. 4. Mapping of fingers to grid cells

tiguous sections of homogeneous elements. In addition, a stabilisation algorithm was implemented to minimise the number of incidental cell changes as shown in Figure 5.

3.2.2 Zoomable Web Browser Interface

With the ever-increasing popularity and use of the World Wide Web, a web-browser interface is arguably more important to a blind user than a desktop or file management system. Our attempts to map web pages into grids similar to our desktop interface proved difficult due to the more free-form nature of interface layouts used. Small items such as radio buttons were forced to occupy an entire cell, and we began to lose the spatial information we were striving to preserve. We therefore opted to discard the grid altogether, and use the native borders of the HTML elements.

Web pages can contain such a wealth of tightly-packed elements, however, that it can take a long time to scan them all and find what you are looking for. To alleviate this problem,

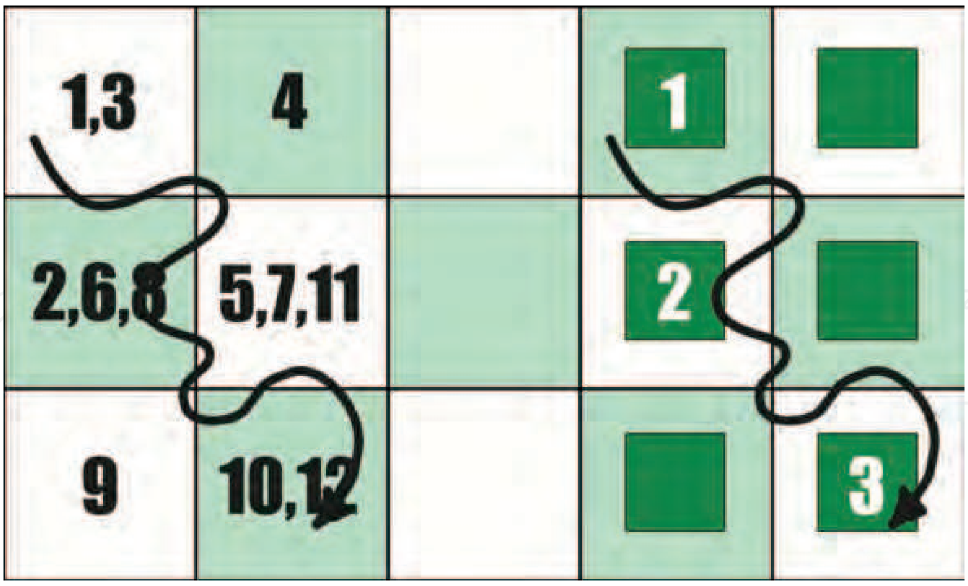

Fig. 5. Gaze travel cell-visiting sequence unstabilised (left) and with stabilisation applied (right)

we took advantage of the natural Document Object Model (DOM) element hierarchy inherent in HTML and "collapsed" appropriate container elements to reduce the complexity of the page. For example, a page containing three bulleted lists containing text and links, and two tables of data might easily contain hundreds of elements. If instead of rendering all of these individually we simply collapse them into the three tables and two lists, the user can much more quickly perceive the layout, and then opt to "zoom" into whichever list or table interests them to perceive the contained elements (see Figures 6(a) and 6(b) for another example).

(a) Raw page (b) Collapsed page

Fig. 6. Example of collapsing a web page for faster perception

Our experimental interface has been developed as an extension for the Mozilla Firefox web browser (Mozilla Foundation, 2009), and uses the BRLTTY (Mielke, 2009) for Braille communication and Orca (GNOME Project, The, 2009) for speech synthesis. It uses JavaScript to analyse the page structure and coordinate gaze-interaction in real-time. Communication with the Braille display (including input polling) is performed via a separate Java application.

3.3 Haptic Output

We have experimented with a number of modes of haptic output during our experimentation, including glove-based electro-tactile stimulation, vibro-tactile actuators, wireless TENS patches and refreshable Braille displays. The following sections discuss the merits of each system.

3.3.1 Electro-Tactile Stimulation

Our initial prototype utilised a simple wired TENS interface as shown in Figure 7. The wires connected the electrodes to our custom-built TENS control unit (not shown). Each electrode delivers a TENS pulse-train of the specified frequency and amplitude (depending on what is being perceived in that region). The voltage and intensity can be varied for each electrode and for each different user. This is necessary given that each user's preferences vary greatly and the sensitivity of different fingers also varies for each individual. This interface proved effective in our experiments and allowed the user's fingers to be free to use the keyboard. However, being physically connected to the TENS unit proved inconvenient for general use.

Fig. 7. Wired TENS system

To eliminate this constraint, we developed wireless TENS patches which communicate with the system via radio transmission. This not only allows the user to walk away from the system without having to detach electrodes, but also enables the electrodes to be placed anywhere on the body such as the arms or torso. A prototype wireless TENS patch can be seen in Figure 8.

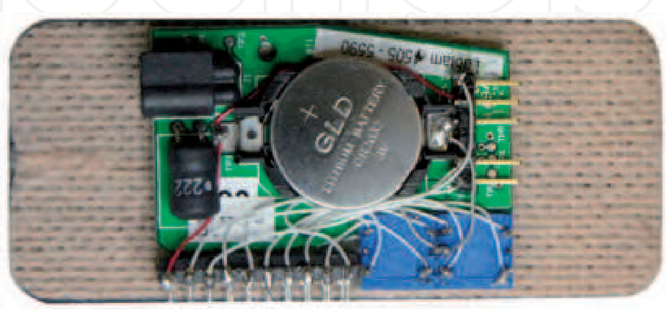

Fig. 8. Wireless TENS Patch

3.3.2 Vibro-Tactile Interface

Although the TENS interface is completely painless, it still requires wireless TENS electrodes to be placed on the skin in a number of places which can be inconvenient. To overcome this problem and trial another mode of haptic communication, we developed a vibro-tactile keyboard interface, as illustrated in Figure 9. This device integrated vibro-tactile actuators, constructed from speakers, which could produce vibration output of the frequency and amplitude specified by the system, analogous to the TENS pulse-train output.

This system has clear advantages over the TENS interface: 1) the user is not “attached” to the interface and can move around as they please, and 2) no TENS electrodes need to be worn and vibro-tactile stimulation is generally more palatable than electro-tactile stimulation despite having a lower bandwidth. Whilst we found this interface capable of delivering a wide range of sensations, the range and differentiability of TENS output was superior. Furthermore, the TENS interface allowed the users to simultaneously perceive and use the keyboard, whilst the vibro-tactile keyboard required movement of the fingers between the actuators and the keys.

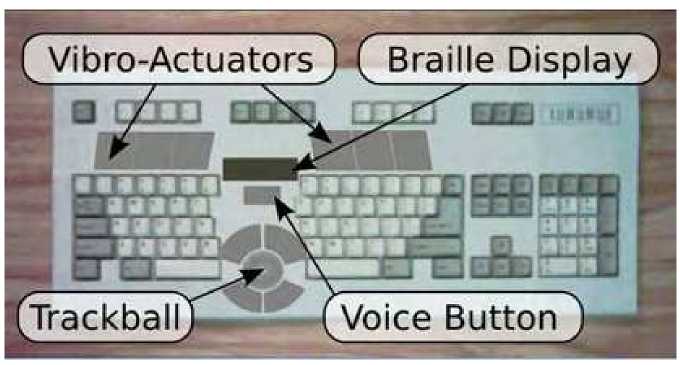

Fig. 9. Vibro-Tactile Keyboard

3.3.3 Refreshable Braille Display

We have also experimented with the use of refreshable Braille displays for haptic perception. Our experimentation revolved mainly around a Papenmeier BRAILLEX EL 40s (Papenmeier, 2009) as seen in Figure 10. It consists of 40 8-dot Braille cells, each with an input button above, a scroll button at either end of the cell array, and an “easy access bar” (joystick-style bar) across the front of the device. We found this device to be quite versatile, and capable of varying the “refresh-rate” up to 25Hz.

A refreshable Braille display can be used in a similar fashion to the TENS and electro-tactile output arrays for providing perception of adjacent elements. Each Braille cell has a theoretical output resolution of 256 differentiable pin combinations. Given that the average user’s finger width occupies two to three Braille cells, multiple adjacent cells can be combined to further increase the per-finger resolution.

Whilst a blind user’s highly tuned haptic senses may be able to differentiate so many different dot-combinations, sighted researchers have significant difficulty doing so without extensive training. For our preliminary experimentation we have therefore adopted simple “glyphs” for fast and intuitive perception. Figure 11 shows some example glyphs representing HTML elements for web page perception.

A further advantage of using a Braille display is the ability to display element details using traditional Braille text. Suitably trained users are able to quickly read Braille text rather than listening to synthetic speech output. Our experiments have shown that using half the display

Fig. 10. Papenmeier BRAILLEX EL 40s Refreshable Braille Display

Fig. 11. Example glyphs – link, text, text

for element-type perception using glyphs and the other half for instantaneous reading of further details of the central element using Braille text is an effective method of quickly scanning web pages and other interfaces.

Fig. 12. Braille Text displaying details of central element

The Papenmeier "easy access bar" has also proven to be a valuable asset for interface navigation. In our prototype browser, vertical motions allow the user to quickly "zoom" in or out of element groups (as described in Section 3.2.2), and horizontal motions allow the display to toggle between "perception mode" and "reading" mode once a element of significance has been discovered.

4. Results

Through this work we have devised and tested a number of human computer interface paradigms capable of enabling the two-dimensional screen interface to be perceived without use of the eyes. These systems involve head-pose tracking for obtaining the gaze position on a virtual screen and various methods of receiving haptic feedback for interpreting screen content at the gaze position.

Our preliminary experimental results have shown that using the head as an interface pointing device is an effective means of selecting screen regions for interpretation and for manipulating screen objects without use of the eyes. When combined with haptic feedback, a blind user is able to perceive the location and approximate dimensions of the virtual screen as well as the approximate locations of objects located on the screen after briefly browsing over the screen area.

The use of haptic signal intensity to perceive window edges and their layer is also possible to a limited extent with the TENS interface. After continued use, users were able to perceive objects on the screen without any use of the eyes, differentiate between files, folders and controls based on their frequency, locate specific items, drag and drop items into open windows. Experienced users were also able to operate pull-down menus and move and resize windows without sight.

The interpretation of screen objects involves devising varying haptic feedback signals for identifying different screen objects. Learning to identify various screen elements based on their haptic feedback proved time consuming on all haptic feedback devices. However, this learning procedure can be facilitated by providing speech or Braille output to identify elements when they are 'gazed' at for a brief period.

As far as screen element interpretation was concerned, haptic feedback via the Braille display surpassed the TENS and vibro-tactile interfaces. This was mainly because the pictorial nature of glyphs used is more intuitive to the inexperienced users. It is also possible to encode more differentiable elements by using two Braille cells per finger.

Preliminary experiments with our haptic web browser also demonstrated promising results. For example, experienced users were given the task of using a search engine to find the answer to a question without sight. They showed that they were able to locate the input form element with ease and enter the search string. They were also able to locate the search results, browse over them and navigate to web pages by clicking on links at the gaze position. They were also able to describe the layout of unfamiliar web pages according to where images, text, links, etc were located.

5. Conclusion

This work presents a novel haptic head-pose tracking computer interface that enables the two-dimensional screen interface to be perceived and accessed without any use of the eyes. Three haptic output paradigms were tested, namely: TENS, vibro-tactile and a refreshable Braille display. All three haptic feedback methods proved effective to varying degrees. The Braille interface provided greater versatility in terms of rapid identification of screen objects. The TENS system provided improved perception of depth (for determining window layers). The vibro-tactile keyboard proved convenient but with limited resolution. Our preliminary experimental results have demonstrated that considerable screen-based interactivity is able to be performed with haptic gaze-tracking systems including point-and-click and drag-and-drop manipulation of screen objects. The use of varying haptic feedback can also allow screen objects at the gaze position to be identified and interpreted. Furthermore, our preliminary experimental results with our haptic web browser demonstrate that this means of interactivity holds potential for improved human computer interactivity for the blind.

6. References

Alexander, S. (1998). Blind Programmers Facing Windows, *Computer World* . Reprinted online by CNN: http://www.cnn.com/TECH/computing/9811/06/blindprog.idg/.

Freedom Scientific (2009a). JAWS keystrokes.
URL: *http://www.freedomscientific.com/training/training-JAWS-keystrokes.htm*

Freedom Scientific (2009b). Job Access With Speech (JAWS).
URL: *http://www.freedomscientific.com/fs_products/software_jaws.asp*

Freitas, D. & Kouroupetroglou, G. (2008). Speech technologies for blind and low vision persons, *Technology and Disability* **20**(2): 135–156.
URL: *http://iospress.metapress.com/content/a4665u784r582844*

GNOME Project, The (2009). Orca.
URL: *http://live.gnome.org/Orca*

Gouzman, R. & Karasin, I. (2004). Tactile interface system for electronic data display system. US Patent 6,762,749.

GW Micro (2009). Window-Eyes.
URL: *http://www.gwmicro.com/*

Hughes, R. G. & Forrest, A. R. (1996). Perceptualisation using a tactile mouse, *Visualization '96. Proceedings.*, pp. 181–188.

Ikei, Y., Wakamatsu, K. & Fukuda, S. (1997). Texture display for tactile sensation, *Advances in human factors/ergonomics* pp. 961–964.

Immersion Corporation (2009). iFeel Mouse.
URL: *http://www.immersion.com/*

Kaczmarek, K. A., Webster, J. G., Bach-y Rita, P. & Tompkins, W. J. (1991). Electrotactile and vibrotactile displays for sensory substitution systems, *Biomedical Engineering, IEEE Transactions on* **38**(1): 1–16.

Kawai, Y. & Tomita, F. (1996). Interactive tactile display system: a support system for the visually disabled to recognize 3d objects, *Assets '96: Proceedings of the second annual ACM conference on Assistive technologies*, ACM, New York, NY, USA, pp. 45–50.

Kieninger, T. (1996). The "growing up" of hyperbraille—an office workspace for blind people, *UIST '96: Proceedings of the 9th annual ACM symposium on User interface software and technology*, ACM, New York, NY, USA, pp. 67–73.

Massie, T. H. & Salisbury, J. K. (1994). The PHANTOM haptic interface: A device for probing virtual objects, *Proceedings of the ASME Winter Annual Meeting, Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, Vol. 55, pp. 295–300.

Maucher, T., Meier, K. & Schemmel, J. (2001). An interactive tactile graphics display, *Signal Processing and its Applications, Sixth International, Symposium on. 2001*, Vol. 1, pp. 190–193 vol.1.

Meers, S. & Ward, K. (2004). A vision system for providing 3d perception of the environment via transcutaneous electro-neural stimulation, *Information Visualisation, 2004. IV 2004. Proceedings. Eighth International Conference on*, pp. 546–552.

Meers, S. & Ward, K. (2005a). A substitute vision system for providing 3d perception and gps navigation via electro-tactile stimulation, *Proceedings of the International Conference on Sensing Technology*.

Meers, S. & Ward, K. (2005b). A vision system for providing the blind with 3d colour perception of the environment, *Proceedings of the Asia-Pacific Workshop on Visual Information Processing*.

Meers, S. & Ward, K. (2008). Head-pose tracking with a time-of-flight camera, *Australasian Conference on Robotics & Automation*.

Meers, S. & Ward, K. (2009). Face recognition using a time-of-flight camera, *Proceedings of the 6th International Conference Computer Graphics, Imaging and Visualization*.

Meers, S., Ward, K. & Piper, I. (2006). Simple, robust and accurate head-pose tracking with a single camera, *The Thirteenth Annual Conference on Mechatronics and Machine Vision in Practice*.

Mielke, D. (2009). BRLTTY.
URL: `http://mielke.cc/brltty/`

Mozilla Foundation (2009). Mozilla Firefox.
URL: `http://www.mozilla.com/firefox/`

Papenmeier (2009). BRAILLEX EL 40s.
URL: `http://www.papenmeier.de/rehatechnik/en/produkte/braillex_el40s.html`

Rotard, M., Taras, C. & Ertl, T. (2008). Tactile web browsing for blind people, *Multimedia Tools and Applications* **37**(1): 53–69.

Siegfried, R. M., Diakoniarakis, D. & Obianyo-Agu, U. (2004). Teaching the Blind to Program Visually, *Proceedings of ISECON 2004*.

Sjöström, C. (2001). Designing haptic computer interfaces for blind people, *Proceedings of ISSPA 2001*, pp. 1–4.

Advances in Haptics

Edited by Mehrdad Hosseini Zadeh

ISBN 978-953-307-093-3

Hard cover, 722 pages

Publisher InTech

Published online 01, April, 2010

Published in print edition April, 2010

Haptic interfaces are divided into two main categories: force feedback and tactile. Force feedback interfaces are used to explore and modify remote/virtual objects in three physical dimensions in applications including computer-aided design, computer-assisted surgery, and computer-aided assembly. Tactile interfaces deal with surface properties such as roughness, smoothness, and temperature. Haptic research is intrinsically multi-disciplinary, incorporating computer science/engineering, control, robotics, psychophysics, and human motor control. By extending the scope of research in haptics, advances can be achieved in existing applications such as computer-aided design (CAD), tele-surgery, rehabilitation, scientific visualization, robot-assisted surgery, authentication, and graphical user interfaces (GUI), to name a few. Advances in Haptics presents a number of recent contributions to the field of haptics. Authors from around the world present the results of their research on various issues in the field of haptics.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Simon Meers and Koren Ward (2010). Head-Tracking Haptic Computer Interface for the Blind, Advances in Haptics, Mehrdad Hosseini Zadeh (Ed.), ISBN: 978-953-307-093-3, InTech, Available from: http://www.intechopen.com/books/advances-in-haptics/head-tracking-haptic-computer-interface-for-the-blind

INTECH

open science | open minds