

## MIT Open Access Articles

*BlindAid: Virtual Environment System For Self-Reliant  
Trip Planning And Orientation And Mobility Training*

The MIT Faculty has made this article openly available. **Please share**  
how this access benefits you. Your story matters.

**Citation:** Schloerb, D.W. et al. "BlindAid: Virtual Environment System for Self-reliant Trip Planning and Orientation and Mobility Training." Haptics Symposium, 25-26 March, Waltham, Massachusetts, USA 2010 IEEE. 2010. 363-370. © 2010 IEEE.

**As Published:** <http://dx.doi.org/10.1109/HAPTIC.2010.5444631>

**Publisher:** Institute of Electrical and Electronics Engineers

**Persistent URL:** <http://hdl.handle.net/1721.1/62165>

**Version:** Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

**Terms of Use:** Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



# BlindAid: Virtual Environment System for Self-Reliant Trip Planning and Orientation and Mobility Training

David W. Schloerb<sup>1</sup>

Touch Lab, Massachusetts  
Institute of Technology

Orly Lahav<sup>2</sup>

Touch Lab, Massachusetts  
Institute of Technology and  
Tel Aviv University

Joseph G. Desloge<sup>3</sup>

Sensimetrics Corporation

Mandayam A. Srinivasan<sup>4</sup>

Touch Lab, Massachusetts  
Institute of Technology

## ABSTRACT

BlindAid is a virtual environment (VE) system that enables blind people to more easily learn about new environments on their own. The system is implemented on a desktop personal computer with a Phantom<sup>®</sup> haptic interface and three-dimensional spatialized audio. In addition to providing blind users with non-visual stimuli similar to what they will depend on in the actual environment, the system is designed to enhance and accelerate the user's understanding of the unknown environment by giving him/her the ability to interact with the VE in ways that are not possible in the real world. The BlindAid system was developed and evaluated as part of an experimental study related to how people who are blind build cognitive maps of their physical environment and how recent advances in VE technology might support orientation and mobility (O&M) training. The current system is able to provide a visual display and record the user's actions in the VE for later review by researchers and O&M instructors. The paper focuses on the technical development of the system with some results from an initial evaluation by four blind volunteers.

**KEYWORDS:** Assistive technology, blind, haptics, human computer interaction, multimodal interface, orientation and mobility, virtual audio, virtual environments

**INDEX TERMS:** K.4.2 [Computers and Society]: Social Issues—Assistive technologies for persons with disabilities; H.5.2 [Information Interfaces and Presentation]: User Interfaces—Auditory (non-speech) feedback and Haptic I/O; K.3.1 [Computers and Education]: Computer Uses in Education—Computer-assisted instruction

## 1 INTRODUCTION

Vision plays the primary role in guiding sighted persons through unknown environments. Unfortunately, for most people who are visually impaired, walking in an unknown environment can be unpleasant, uncomfortable, and unsafe. We developed the BlindAid system, a non-visual virtual environment (VE), to provide a means for people who are blind to safely explore and learn about new spaces in advance on their own. A potential future application might be to use such a system to download virtual maps from the internet similar to the way sighted people use Mapquest<sup>®</sup> or Google<sup>™</sup> Maps. Although our initial work has

focused on people with total blindness, people with low vision might also benefit from the technology with the development and integration of a suitable low-vision visual display.

The current system is also a research tool and a learning tool for enhancing orientation and mobility (O&M) training. We are using the system to study how people who are blind build cognitive maps of their physical environments as well as evaluating its potential within an O&M curriculum.

This paper discusses the technical development of the BlindAid system. After a brief background discussion, the system design is presented in Section 3. Then some key design issues are discussed (Section 4) and related results from an initial evaluation by blind volunteers are summarized (Section 5).

## 2 BACKGROUND

Haptic devices that enable manual interactions with a VE have been developed for a wide range of applications, including medical, industrial, educational, scientific, and entertainment [1]. The most successful haptic devices to date simulate touching the virtual world through some type of handle, akin to probing the real world with a stick.

Spatialized audio systems give the listener the impression of being surrounded by sound sources in a three-dimensional (3D) environment. Such systems require sophisticated signal processing, taking into account the possibly time-varying location/orientation of the listener with respect to the sound sources plus the shape of the person's head and ears. In addition to VE applications, these audio displays are useful for warning systems and speech communication [2, 3].

Both haptic and audio VE interfaces have been investigated for O&M applications for people who are blind. The BlindAid system described in this paper evolved from the prior work of one of the authors, Lahav [4]. She developed a 2D VE system and used it to study navigation and cognitive mapping by people who are blind. The system simulated a single room with non-spatialized sounds, with haptic feedback provided by a Sidewinder Force Feedback Pro Joystick (Microsoft). Another system that used consumer-quality haptics was developed by others for displaying non-visual geographic maps [5]. The latter incorporated a low-cost spatialized sound system that enabled listeners to distinguish between sources right and left, with some separation between front and back. The initial results with these systems were promising.

Research-quality haptic devices (like the one used with BlindAid) have been used to investigate various design components of the proposed application. Studies of how texture and shape information are communicated through a haptic device suggested that this technology was potentially useful for blind users [6]. Another study investigating how haptic devices might enable improved computer interfaces for the visually impaired, concluded with the suggestion that a "very useful application for blind people is to create haptic maps and models of public spaces" [7]. In addition, preliminary experiments related to haptic

---

1 schloerb@mit.edu

2 lahavo@mit.edu

3 desloge@sens.com

4 srini@mit.edu

zooming and scrolling in a non-visual VE have been performed [8].

Primarily-audio VE systems (without haptics) have also been developed for blind users [9]. An advantage of audio-based systems is that they may be incorporated into mobile devices for navigation in situ [10, 11].

Perhaps the non-visual VE system with the most realistic simulation to date is the HOMERE system [12]. The system simulates a full-size white cane and audio is displayed via four surround-sound speakers. In addition, heat lamps are mounted in a ring above the user to simulate thermal directional information from the sun. A major limitation with the reported system [ibid.] is that motion within the VE must follow a predefined path. The system is also relatively large, about the size of an office cubicle.

### 3 SYSTEM DESCRIPTION

The BlindAid system provides VE maps for people who are blind and consists of application software running on a personal computer equipped with a haptic device and stereo headphones. The haptic device used in our experiments was a Desktop Phantom<sup>®</sup> (Sensable Technologies) and we used Sennheiser HD580 headphones. Different VEs, corresponding to virtual maps of different environments to be explored, are saved as text files that can be opened by the application. Figure 1 presents a photograph of the system.



Figure 1. Photograph of the BlindAid system in use. The user hears spatialized sounds, as if physically standing in the simulated space, while touching a small scale-model of the space (~10 cm across) using the Phantom. The user may input commands via help keys on the computer keypad. A video display allows sighted O&M instructors and researchers to monitor the user's progress.

The BlindAid software is a multi-threaded C++ application running under Windows XP. It uses OpenGL to render graphics, OpenHaptics<sup>™</sup> (SensAble Technologies) to render haptic forces to the user via the Phantom, and a custom audio library to display 3D spatialized sounds via the headphones. A standard Windows graphical user interface is used to start/exit the application, select/open/close VE maps, and for other tasks associated with research, training, and VE development.

The Phantom haptic device allows the blind user to interact manually with a scale-model of the currently open VE. The haptic device has two primary functions, (1) it controls the motion of the *avatar* (a representation of the user) in the VE and (2) it provides force feedback to the user that gives cues about the space similar to those generated by a white cane. In operation, the user holds a stylus, like a pen, which has its tip connected to the rest of the

Phantom mechanism through a gimbal. As the user moves the Phantom stylus, the mechanism tracks the position of the stylus tip in the physical workspace and the system continuously updates the position of a corresponding virtual point in the VE. The corresponding point in the VE is the user's avatar position, which the user may also think of as the location of the tip of a white cane that he/she is holding. When the avatar comes in 'contact' with an object in the VE, motors in the Phantom cause it to push back against the tip of the stylus as if the tip had come in contact with a real object.

In the current implementation, the virtual space is bounded by two horizontal planes: *floor* and *ceiling*. VE developers create virtual maps by defining static virtual objects at fixed locations within this horizontal space. All objects extend uniformly from floor to ceiling and are represented by a small set of standard object types comprised of simple geometric shapes with various properties. For example, the *wall* type is a vertical rectangular plane that is used to represent walls. All walls have the same graphic and haptic properties, but individual walls may be associated with unique identifying sounds. Another example, the *rectangle* type (a right rectangular prism, or box shape, with tops and bottoms parallel to the floor), may be used to represent a wide range of real world objects (e.g., tables, chairs). Each *rectangle* object has its own individual graphic, audio, and haptic properties. A single object in the VE may also be represented by more than one instance of a given virtual object type and possibly as a combination of instances of multiple types. Other standard object types include: *windows*, *private doors*, *public doors*, and arbitrary *vertical boundaries* (all of the preceding are vertical planes similar to walls, but with potentially different properties), *cylinders* (similar to the *rectangle* type), and *ground textures* which define regions of texture on the floor (see Section 3.4).

#### 3.1 Graphics

Simple graphics are presented on the computer monitor (an orthographic view of the *virtual workspace* from above; see Section 3.3), intended for VE developers, researchers, and O&M instructors. The graphics appear in the client area of the application's main window, which is constrained to have the same aspect ratio as the Phantom's usable workspace when the window is resized.

#### 3.2 Audio

3D audio allows the user to hear the direction and distance of sounds in the VE as if the person were standing at the location of the avatar at full scale. The 3D audio processing used in the BlindAid system employs the head-related transfer functions (HRTFs) of the Knowles Electronic Manikin for Acoustic Research [13] as measured by Gardner and Martin [14]. A simple, single-channel sound signal is converted into a spatialized, 3D-audio, stereo sound signal as follows. First, the signal is scaled so that the level corresponds to the level when listening at a distance of 1 meter. Second, the source angular position (i.e., azimuth and elevation) is determined relative to the avatar's head and the signal is processed through the pair of HRTFs that describe anechoic source-to-left-and-right-ear propagation for sources located at the specified angular position at a distance of 1 meter. Third, the resulting stereo signal is scaled to reflect source distance based on the 1-meter reference distance and the virtual source distance from the avatar's head, but with the overall sound level capped to prevent excessive amplification of nearby sources. Fourth, the signal may be filtered to indicate the number of walls, if any, that are located between the source and the avatar. We used the same successively-sharp lowpass filters for this purpose in all of our tests. Finally, the signal is corrected for headphone response and presented to the user.

By default, the BlindAid 3D-audio processing assumes that the avatar's head orientation is fixed and facing forward in the VE, corresponding to the nominal orientation of the user's head in the physical workspace of the Phantom. A second audio display mode is also available that allows the user to dynamically control the azimuth-plane orientation of the avatar 'head' by rotating the Phantom stylus about its long axis. This mode, which mimics turning the head side-to-side in the real world, is intended to provide supplemental 'head-movement' cues that can improve sound localization by disambiguating sources located symmetrically front and back of the listener. Finally, a third mode disables the angular localization cues by not processing the sound for angular location (but still scaling for distance) and presenting the sound monaurally.

The display sound clips are pre-recorded by the VE developer and linked to various features as needed when the VE is created. There are three general 3D sounds:

- *Contact* sounds are generated when the user avatar comes in contact with objects in the VE in order to provide the user with information about the objects. The location of the contact sound is defined behind the object surface so that the user hears the sound coming from the appropriate direction in stereo mode. Typically, the contact sound is a short iconic sound or *earcon*. For example, normally, a brief 'tick' sound might be produced when a wall is contacted, with the same sound common to all walls. Similarly, all rectangle objects that represent chairs might cause the same 'ding' when touched. Short contact sounds might also be a short verbal description such as "elevator." Alternatively, pressing a specific key on the keyboard while touching an object substitutes a longer verbal description of the individual object if one is specified.

- *Background* sounds are associated with a specific area within the VE and serve to provide the user with cues about his/her location, similar to ambient sounds in the real world. A background sound source is defined to be at a particular point within its specified region and it plays continuously while the avatar is within the region.

- *Landmark* sounds play in response to a key press and serve as audio beacons. Up to 3 landmarks may be pre-defined within a virtual map at fixed locations by the VE developer in order to mark important places. The location of 6 other landmarks may be modified by users during their explorations to help them find their way back to points of interest.

In addition to the general 3D sounds above, the BlindAid system also implements a reverberant tap feature in an attempt to give users the ability to sense the proximity of nearby structures in the VE, similar to the way some people are able to perceive their surroundings in the physical world by listening to the tap of a cane. Specifically, simulated tap sounds and corresponding echoes are generated when the user presses a specific key on the keyboard. However, we did not evaluate the reverberant tap feature in our experiments with blind subjects.

### 3.3 Haptics and Control

Haptics is inherently bidirectional and, as described previously, the Phantom is used for both sensing and control in the VE. Fourteen *help keys* on the computer keyboard (the numeric keypad plus the arrow keys) also provide a means for the user to control the system while interacting with the VE. See Table 1 for a complete listing of the help keys.

The *virtual workspace* is a standard feature of the 3D haptic interface, common to all VEs that may be simulated by the BlindAid system. The virtual workspace is a rectangular box within the VE that corresponds to the usable physical workspace of the Phantom. The virtual workspace has virtual sides, ceiling, and a floor to keep the user inside the Phantom's usable region and these boundaries have an identifiable feel which informs the

user when he/she has reached the end of the workspace. The ceiling and floor of the workspace correspond to the ceiling and floor of the VE. The side boundaries, however, have no unique relationship to points in the VE, which is effectively infinite in all horizontal directions.

Table 1. Help keys.

Key	Name	Function
0	Restart	Returns user to the start of the virtual map.
1	Replay	Not implemented. See Section 4.1.3.
2	Pause/Start	First press causes the Phantom to hold its position; second press releases it.
3	Undo	Undoes the affect of the last help key press.
4	Zoom In	Changes the horizontal haptic scale to give more detail.
5	Landmark	Pressed in sequence with other keys to play or modify audio landmarks.
6	Zoom Out	Changes the horizontal haptic scale to give the 'big picture.'
7	Tap	Generates a reverberant tap sound at the avatar location.
8	Head Rotation	While pressed, the user may control the orientation of the avatar's head by rotating the Phantom stylus about its long axis.
9	Long Sound	While pressed, alternate "long" descriptions are played when the user contacts objects in the VE.
→	Move Right	Moves the workspace to the right in the VE. Left, Forward, Back are similar.

Dimensions in the VE are specified in meters and, when the VE represents a real space, the size and spacing of objects in the VE correspond to the actual dimensions in the real space. Because the range of motion of the Phantom DeskTop model we used is only about 10 cm, this normally requires that the size of objects displayed to the user by the haptic device be scaled-down from the VE to the physical workspace.

While exploring an environment, the user may control both the horizontal scale and the location of the workspace within the VE. For example, by pressing the Zoom Out help key, the VE is scaled down so that a larger region fits into the workspace, allowing the user to experience the 'big picture.' Alternatively, the user may *zoom in* to feel a smaller piece of the VE in more detail. The allowable horizontal scales, or discrete zoom levels, are specified by the VE developer in advance for each VE. In addition, every object within the VE is assigned a maximum zoom-out level. This allows the developer to control the amount of detail that may be felt at the different zoom levels because smaller objects can be made to disappear as the user zooms out. The vertical haptic scale is not an issue because in the current version all virtual objects extend uniformly from floor to ceiling.

There are two methods for moving the workspace within the VE in order to explore beyond the horizontal confines of the workspace. The first method involves pressing one of the arrow help keys. Each arrow key press shifts the workspace 1/2 its width in the given direction. When this happens, the Phantom gently and simultaneously, moves the users hand an equal distance in the opposite direction such that the avatar position in the VE remains unchanged. The workspace stops moving if the Phantom reaches the opposite physical boundary before the nominal move is complete. The second method involves only the Phantom. With this method--for example, upon reaching a side of the workspace--the user presses a button on the stylus causing the user avatar position to be fixed in the VE. Then, similar to the way one repositions a computer mouse upon reaching the edge of the

mouse pad, the user moves back from the physical workspace boundary while holding the stylus button, causing the virtual workspace to advance in the opposite horizontal direction.

On the sensory side, the reaction forces felt by the user in response to his/her manipulation of the Phantom can simulate contact with a wide range of object geometries and mechanical surface properties. Four general parameters are provided by the OpenHaptics library to control surface properties: stiffness, damping, static friction, and dynamic friction [15, Table B-13]. A fifth haptic parameter, popthrough [ibid.], is always 1 except in the case of *public doors* (one of the standard object types), in which case the parameter is 0.1 so that the user may pass through the door by applying a modest force.

### 3.4 Texture

The haptic device is also able to simulate texture. The BlindAid system simulates textures using the 2D method developed by Minsky [16]. In this technique, the feeling of sliding over a vertical displacement on a surface (i.e., displacement normal to the surface) is simulated by applying a horizontal force to the hand (i.e., force tangential to the surface). The tangential texture force is equal to the tangential force that would be needed to balance the normal force applied by the user, given the slope of the virtual texture at that point and assuming no friction.

Although sophisticated methods of haptic texture mapping [17] can be implemented in the future, in our current implementation, it was sufficient to define textures for a given plane surface by specifying one cycle of a repeating series of uniform vertical ridges. Figure 2 illustrates the three types of textures ridges that may be specified and Figure 3 shows the dimensions that are used to define the ridge parameters. The ridge slopes used to calculate the texture force are pre-calculated from the specified parameters at 128 points on the wave cycle to speed up rendering. The texture dimensions are defined in the physical workspace of the Phantom and are, thus, the same for all haptic zoom levels. Table 2 lists and describes the individual parameters.

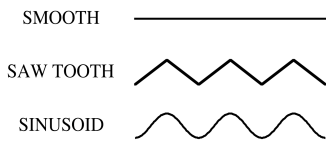


Figure 2. Texture ridge types.

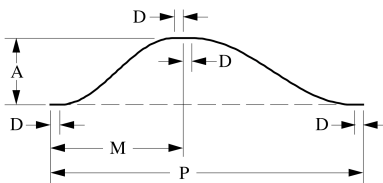


Figure 3. Texture ridge dimensions. The figure shows the Sinusoid case. The Saw Tooth case is the same except that the sinusoidal segments are replaced by straight lines. The horizontal axis in the figure corresponds to the texture reference frame axis ( $X_0$  or  $X_1$ ; see Figure 4) and the vertical axis is normal to the surface (up in Figure 3 corresponds to out of the paper in Figure 4). Note, however, that there is no displacement normal to the virtual surface associated with the texture. Instead, as discussed in the text, the slope of the virtual ridge is used to calculate horizontal forces that simulate vertical ridges.

Table 2. Texture ridge parameters.

Parameter	Description (see Figure 3)
Type	Smooth, Saw Tooth, or Sinusoid. In the case of a smooth texture, the other ridge parameters are irrelevant.
Amplitude	Peak to peak amplitude or ridge height, normal to the surface, $A$ (mm).
Period	Period or wavelength, distance between ridges on the surface, $P$ (mm).
Up Ratio	Ratio of up slope region to total period ( $M/P$ ).
Deadband Ratio	Ratio of deadband regions to total period ( $4D/P$ ). The four equal length ( $D$ ) deadband regions are flat (zero slope) segments of the waveform.

In addition to the ridge parameters, two orthogonal sets of virtual texture ridges may be defined on a given surface in the VE and the texture reference frame may be rotated ( $\pm 45^\circ$ ) relative to the reference frame of the surface (Figure 4). The full specification of texture on a surface thus involves two sets of ridge parameters corresponding to the  $X_0$  and  $X_1$  texture directions plus the rotation angle ( $\theta$ ). Texture reference frames are defined relative to the surfaces to which they are attached and the surfaces themselves may have any position and orientation within the VE.

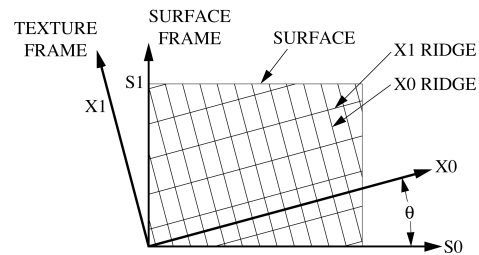


Figure 4. Texture reference frame. The lines representing the ridges in the figure correspond to the locations of the ridge peaks.

### 3.5 Operational Modes

In addition to the primary *User* mode, the system also has *Edit* and *Test* modes for VE developers, researchers, and O&M instructors. The Edit mode allows developers to create and modify VEs. A key feature of the Edit mode is the program's ability to import AutoCad (dxf) 2D drawing files to aid in the creation of new virtual maps. This makes it possible, for example, to import a floor plan so that the locations of all of the walls of a building are accurately replicated in the VE.

The Test mode allows researchers and O&M instructors to record and review a user's exploration activities in the VE. During an experiment or a learning session, the computer records the avatar's position and orientation within the VE from moment to moment in a text file, along with any command actions. Subsequently, the data file may be viewed directly with a text editor or replayed by the system like a video recording. Figure 5 presents the system's replay display.

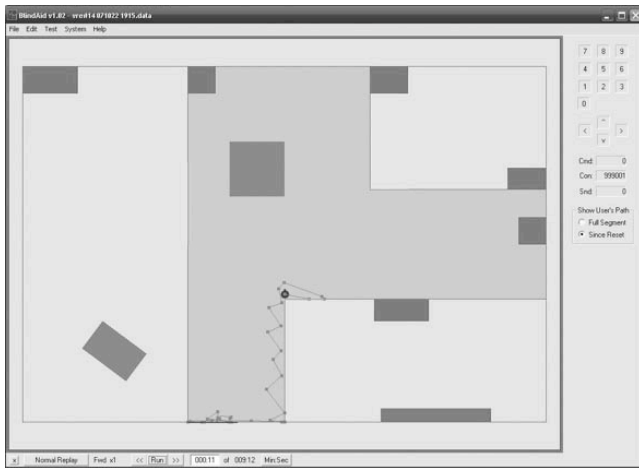


Figure 5. Replay display. The central part of the display is identical to what a researcher/instructor would see during a test/learning session, except the user's path may also be displayed (dots interconnected by lines near bottom center). Also, the control/indicator parts of the window along the right and bottom are only visible during replay. Controls along the bottom allow the researcher/instructor to start, stop, reverse, jump to a specific time sample, or change the speed of the playback. The indicators at the upper right change color corresponding to when the user pressed one of the help keys during the test. Numeric codes are also listed for every user command along with codes for contact with objects and sounds played. During playback, the researcher/instructor may also listen to the sounds that were heard by the user.

## 4 DESIGN ISSUES

There are a number of interesting design issues associated with the BlindAid system. Following are the key issues that we faced.

### 4.1 Non-visual VE Issues

A number of display features are possible with visual multimodal VE systems that include haptics, which become an issue if the visual channel is not present. For example, consider the case of moving the virtual workspace from one location in the VE to another. See Section 3.3 for the definitions of *virtual workspace* and other terms used in this discussion.

Moving the workspace is easy with a visual display: you just change the image on the display. This works because vision is primarily a sensory channel that does not affect the surrounding environment. The same is true for audio displays. Also, a single point of view encompasses an entire region of the workspace for both vision and hearing, making it easier to give the user a frame of reference and continuity when moving from one location to the next. Indeed, with audio there is no limitation like the finite haptic workspace--the user is able to hear objects at any point within the VE. In contrast, haptics involves both sensory input and control output and the instantaneous haptic interaction occurs at only one location out of many within the workspace (in our case, a single point). In other words, vision and hearing are global sensory channels while haptics is bi-directional and local. Further, the mapping of the haptic interaction point in the virtual workspace (the avatar location in the VE) must match the corresponding mapping between the Phantom's physical workspace and the tip of the Phantom stylus. And, you can't change one without affecting the other.

What this means in our example is that if you move the workspace, you must either (1) adjust the location of the avatar in the VE to match the Phantom stylus position or (2) physically move the stylus to put (or keep) the avatar in the right place.

Either approach may be disorienting if the user lacks a reference for the new location and both approaches have pros and cons. The first approach can land the avatar inside of solid objects--something that is not typically a problem when vision is present because the user can see when and where he/she emerges from the object (assuming that it is haptically transparent from the inside). The second approach avoids the problem, but is harder to implement because it involves physically moving the user.

Following are specific non-visual VE issues and how we handled them.

#### 4.1.1 Moving the Virtual Workspace

Corresponding to the example just described, we implemented two methods for moving the workspace: the arrow key method and the Phantom stylus method. In both cases, the avatar stays at the same place within the VE during the move so that the user knows where he/she is at all times. This also ensures that the avatar stays at a physically realizable location within the VE. As described previously, with the arrow key method the Phantom moves the user's hand equal and opposite to the movement of the workspace, where as, with the Phantom stylus method it is the movement of the user's hand that causes the virtual workspace to move in the opposite direction.

#### 4.1.2 Haptic Zoom

Changing the linear scale of the virtual workspace with respect to the physical workspace, or zooming is also an issue with a non-visual VE. The first consideration is selecting the *zero point* that does not move when the space is scaled. We opted to use the avatar location as the zero point so the user would stay in the same place during the zoom. A potential drawback of this approach is that, if the user moves the avatar after zooming in one direction (in or out) and then zooms back, the virtual workspace ends up in a different place. Because moving the avatar is an essential part of exploring the haptic space, changing the workspace location is almost an inevitable part of haptic zooming with a non-visual VE.

A second potential haptic zoom-related issue for non-visual VEs is how easy it might be for the user to keep track of the change in scale. Our initial approach to this issue was simply to limit the number of zoom levels to three in our experiments (the system allows up to 14 zoom levels). The system also 'beeps' to indicate that the key press has no effect if the user tries to zoom beyond the minimum or maximum haptic zoom levels.

#### 4.1.3 Replay, Restart, and Undo

One feature that we originally hoped to include in the system was a Replay help key. The proposed idea involved having the Phantom stylus guide the user's hand back over a previously explored path when he/she pressed the help key, returning the user to where he/she began. A related idea is a more general guide agent that would lead the user to points of interest, like a human guide introducing the user to a new space. The goal in either case is to help train the user how to navigate in the space. Unfortunately, because haptics is bi-directional, the experience of being led by the hand is quite different from actively exploring with one's hand. The optimal technique for achieving the desired goal of simulating an expert guide agent needs further research. We did, however, implement Restart and Undo help keys.

When pressed, the Restart key transports the user to the developer-defined starting point on the virtual map. This involves moving and zooming the virtual workspace to match the desired starting location and scale within the VE, plus having the Phantom move the user's hand to the point in the physical workspace that matches the starting point in the VE (see preceding discussion in Section 4.1).

While the Restart key does not help the user to learn his/her way to the starting point (as a guide agent might), it does at least take him/her to a known location. A straightforward extension of this idea would be to provide a help key sequence that would transport the user to any of the landmark locations (as an alternative to simply hearing the landmark sound; see Section 3.2). However, we did not implement this feature.

Similar to the undo command in many computer programs, the BlindAid Undo command undoes the affect of the immediately preceding help key press. Undo was implemented for all commands for which it was appropriate (except Undo itself, i.e., Redo was not implemented). For example, undoing a Restart command returns the user to the point in the VE where the user was when he/she pressed the Restart key. This involves both moving the virtual workspace and avatar in the VE and moving the Phantom (user's hand) in the physical workspace as needed to return all of them to exactly the same virtual or physical place. This happens regardless of any intervening movement of the Phantom as long as no other undoable help key was pressed. Undoing the zoom and arrow key commands works in the same way. In the case of a new landmark location, the previous landmark coordinates are simply restored.

#### 4.1.4 Global Spatial Cues

Because hearing is a global sense, with perception over a region of space, it can and should be used to mitigate many non-visual VE issues associated with the local nature of haptics. For example, the BlindAid system has background and landmark sounds to help the user stay oriented. We also use a unique sound to inform the user when he/she has passed through a public door, an event that we discovered early on is hard to detect without this cue (at least for a sighted user with eyes closed). We did not develop special haptic cues to enable a more global sense of touch--like a gentle vibration when the avatar comes near a surface--although this seems like an interesting idea to explore in the future.

#### 4.2 Audio Modes and Audio Display Transformation

In a 'real' environment, one in which a person is physically present, the physical stimuli associated with perceived direction and distance are consistent across all sensory modalities. Hence, for example, if you were to reach out to touch a bell, you would normally expect to feel it, see it, and hear it ring at the same location. There is no such restriction in a VE, however, and the relationship between the VE and the audio display was an open issue in the design of the BlindAid system. Transforming audio and haptics in the same way does not make much sense because it would tend to make everything sound as if it was in a small region in front of the user (specifically, any sound source located in the virtual workspace would be heard as if it were located inside the Phantom's physical workspace). Instead, our intuition was that the user would best be able to understand spatial audio cues, heard as if he/she were actually standing in the VE at full scale, even though he/she is simultaneously touching a small scale model of the VE through the Phantom. This is the only audio display transformation that we implemented.

Interpreting the audio cues is easiest with the monaural mode (which lacks directional information and, hence, there is less information to take in), harder with stereo, and hardest with the "stereo with Phantom rotation" mode (which requires the user to mentally translate rotation of the Phantom stylus to rotation of his or her own head in order to interpret the direction of the sound).

#### 4.3 Representation vs. Reality

Maps are inherently simplified representations of the world and we designed the BlindAid system to display a greatly simplified

virtual world. Real-world objects are represented in the VE by simple geometric shapes with associated sounds and touch sensations to help the user distinguish one object from another. The haptic space is 3D, but it might be better described as 2-1/2D because the entire VE is a single horizontal region with all objects extending from floor to ceiling. The virtual audio space is also essentially 2D because the complex spectral cues that assist in elevation localization are typically very difficult to achieve with the type of system we used which employs generic HRTFs [18].

Our 3D audio design choice was to accept the limitation on elevation localization because the measurement of individual HRTFs is a time-consuming process that may not be feasible in many circumstances. Moreover, even with personalized HRTFs, the display system is very sensitive to factors such as placement of the headphones on the head—a single re-positioning of the headphones can reduce the realism of the localization cues [19]. This problem may be overcome in the future because recent work suggests a method for modifying generic, non-personalized HRTFs to improve 3D source localization performance [20].

Beyond working within the limits of the technology, the challenge is to find ways to represent information so that it might be communicated to the user better. By "better" we mean faster or more effective in terms of training than would be possible in a real environment. Other performance metrics include cost or ease of use, comparing different VE systems or alternate training methods.

### 5 INITIAL EVALUATION

This section summarizes experiments in which the BlindAid system was evaluated in terms of usability and user preference. These tests were part of two larger experimental studies that investigated spatial cognitive mapping and learning strategies in the VE (Study 1) and the transfer of spatial knowledge from a VE to the real world (Study 2).

Four adult human subjects who were totally blind volunteered to take part in the tests in accordance with an approved human subject protocol. The subjects, 3 males and 1 female, 41-53 years old, answered a questionnaire on their initial O&M skills and there were no apparent differences among them in the context of the test.

The four subjects learned how to operate the system as part of the studies, exploring 17 different VEs of increasing complexity. The subjects were tested after each exploration in terms of how well they had been able to learn the unknown space. Specifically, in Study 1--after some initial tests that focused on training (VEs #1 and #2) and user preference with regards to haptic feedback (VEs #3 to #8)--the subjects were asked to describe VEs #9 to #13 verbally and to build physical models of them using a modeling kit designed for the tests. In Study 2, in which the VEs (#14 to #17) represented real physical spaces, the subjects were tested in their ability to navigate in the real spaces after training only in the VEs. The experimental protocol in both studies also involved a standard set of verbal questions to elicit statements from the subjects about their preferences. In addition, the experimenter observed which features of the system were used and how they were used. The sessions were video-recorded and the subjects' explorations were also recorded by the BlindAid system to aid the experimenter in this task.

In general, all of the subjects were able to learn how to operate the system and use it to learn about unknown spaces. They also demonstrated the ability to transfer and apply spatial knowledge gained in VEs to navigation in real spaces. Following is a brief summary of the results from the tests related to the design issues discussed in the previous section (4).

## 5.1 Non-visual VE Issues

### 5.1.1 Moving the Virtual Workspace

Each subject was trained to use both methods for moving the workspace in two separate VEs (#12 and #13). At the end of the two sessions, each of the four subjects chose to use the Phantom stylus method for further tests. The subjects reported that the Phantom stylus method was much more intuitive than the arrow key method and it gave them more control over movements.

### 5.1.2 Haptic Zoom

All four subjects were able to operate the zoom feature of the system. An interesting point is that the zoom concept, which is a visual idea, was completely foreign to the subjects initially. However, once they understood it, they were able to use the zoom feature independently with no observed problems.

### 5.1.3 Replay, Restart, and Undo

As discussed previously, the Replay help key was never implemented. When asked for suggestions about how the system might be improved, two subjects independently came up with our idea of extending the functionality of the Restart help key to the landmarks (potentially a replay-like feature). It should be noted that all of the subjects spontaneously used the Restart and Pause/Start help keys in the more complex environments, suggesting that these are useful features. While none of the subjects ever used undo in the tests, we believe that it is still too soon to discount the usefulness of that feature.

### 5.1.4 Global Spatial Cues

The subjects were able to use background and landmarks sounds to help them stay oriented in the VE, although one subject preferred to use the monaural audio mode that substantially eliminates the benefit of these features (see Section 5.2). Because of that result, we feel that future work is needed to develop new audio and haptic features that will enable better global perception in non-visual VEs. Our current research has only begun to explore this area.

## 5.2 Audio Modes and Audio Display Transformation

Three out of the four subjects adapted easily to, and preferred the stereo audio mode. They reported that stereo audio helped them to determine the direction of objects and to stay oriented in the VE. This result suggests that the disparity between haptic and audio display transformations was not a problem for these subjects. The 4th subject was also able to hear directional information in stereo mode, but preferred the monaural audio mode to reduce the amount of information he had to keep track of while using the system. This suggests that the 4th subject had some difficulty adapting to the transformation disparity. However, the primary issue for this subject was information overload, such that a mixed monaural/stereo mode that allowed him to hear only certain sounds in stereo (like landmarks) might have been acceptable. The same subject found the stereo with Phantom rotation mode “confusing” and, indeed, none of the subjects liked the rotation mode because of the extra effort required.

## 5.3 Representation vs. Reality

As a general rule, the subjects preferred their interactions with the VE to be simple and found that the VE presented a useful representation of the world. All of the subjects said that they preferred a small number of different haptic properties to be used to represent different structures and objects within the environment. For example, they preferred having different general haptic properties for walls, windows, and doors. They liked a hard/soft distinction between objects (e.g., hard for a table, soft

for a sofa). They preferred that a unique rigid texture be used to designate areas/objects that are a safety issue (e.g., stairs). One reason for this is that they tended to find the small set of textures that we tested to be unpleasant.

The subjects used the Phantom primarily for understanding spatial relationships and the audio to gain more information about objects once they had been found. As one subject observed, as soon as he has located something, he does not “care what it feels like anymore.”

The subjects also liked having objects identified by earcons as they explored. While their preference was for the meaning of the short contact sounds to be self-evident (without needing to be learned), even if the earcon had to be learned, they felt that the alternative of repeatedly hearing a long description every time a common object was touched was tedious. Even a single word (e.g., “wall”), takes much longer than a person is willing to wait before moving on during an exploration with the Phantom. On the other hand, the subjects made repeated spontaneous use of the long contact sounds to gather more information as needed about specific structures and objects within the VE.

Finally, while we do not presume to answer the question about whether a more realistic ‘immersive’ VE might make it easier or harder for the user to learn about an unknown environment, our results do suggest that a relatively simple system can be effective. The key to success is making judicious choices in the design, appropriate to the available technology.

## REFERENCES

- [1] S. J. Biggs and M. A. Srinivasan, "Haptic Interfaces," in *Handbook of Virtual Environments: Design, Implementation, and Applications*, K. M. Stanney, Ed. Mahwah, NJ: Lawrence Erlbaum Associates, 2002, pp. 93-116.
- [2] D. R. Begault and M. T. Pittman, "Three-Dimensional Audio Versus Head-Down Traffic Alert and Collision Avoidance System Displays," *The International Journal of Aviation Psychology*, vol. 6, pp. 79-93, 1996.
- [3] J. A. MacDonand, J. D. Balakrishnan, M. D. Orosz, and W. J. Karplus, "Intelligibility of speech in a virtual 3-D environment," *Human Factors*, vol. 44, 2002.
- [4] O. Lahav and D. Mioduser, "Exploration of Unknown Spaces by People who are Blind, Using a Multisensory Virtual Environment (MVE)," *Journal of Special Education Technology*, vol. 19, 2004.
- [5] P. Parente and G. Bishop, "BATS: The blind audio tactile mapping system," presented at 41st Annual Southeast Regional Conference (ACM-SE 41), Savannah, GA, 2003.
- [6] G. Jansson, H. Petrie, C. Colwell, D. Kornbrot, J. Fänger, H. König, K. Billberger, A. Hardwick, and S. Furner, "Haptic Virtual Environments for Blind People: Exploratory Experiments with Two Devices," *International Journal of Virtual Reality*, vol. 4, 2000.
- [7] C. Sjöström and K. Rasmus-Gröhn, "The sense of touch provides new computer interaction techniques for disabled people," *Technology and Disability*, vol. 10, pp. 45-52, 1999.
- [8] C. Magnuson and K. Rasmus-Gröhn, "Non-visual Zoom and Scrolling Operations in a Virtual Haptic Environment," presented at Eurohaptics, Dublin, Ireland, 2003.
- [9] J. Sánchez and L. Jorquera, "Interactive virtual environments for blind children: Usability and cognition," presented at Iberoamerican Congress on Informatics and Education, RIBIE 2000, Viña del Mar, Chile, 2000.
- [10] J. M. Loomis, R. G. Golledge, R. L. Klatzky, and J. R. Marston, "Assisting wayfinding in visually impaired travelers," in *Applied Spatial Cognition: From research to cognitive technology*, G. L. Allen, Ed. Mahwah NJ: Lawrence Erlbaum Associates, 2007.



- [11] S. Landau, W. Wiener, K. Naghshineh, and E. Giusti, "Creating Accessible Science Museums With User-Activated Environmental Audio Beacons (Ping!)," *Assistive Technology*, vol. 17, pp. 133–143, 2005.
- [12] A. Lecuyer, P. Mobuchon, C. Megard, J. Perret, C. Andriot, and J. P. Colinot, "HOMERE: a multimodal system for visually impaired people to explore virtual environments," presented at IEEE Virtual Reality, Los Angeles, CA, USA, 2003.
- [13] M. D. Burkhard and R. M. Sachs, "Anthropometric Manikin for Acoustic Research," *J. Acoust. Soc. Am.*, vol. 58, 1975.
- [14] B. Gardner and K. Martin, "HRTF Measurements of a KEMAR Dummy-Head Microphone," MIT Media Laboratory 280, May 1994.
- [15] Sensable Technologies, "Open Haptics™ Toolkit, version 2.0, API Reference," 2005.
- [16] M. D. R. Minsky, "Computational haptics : the Sandpaper system for synthesizing texture for a force-feedback display," in *Architecture, Program in Media Arts & Sciences*. Cambridge: Massachusetts Institute of Technology, 1995, pp. 217.
- [17] Basdogan C and Srinivasan MA, *Haptic Rendering in Virtual Environments*, *Virtual Environment Handbook*, Ed: K. M. Stanney, Lawrence Erlbaum Associates, Ch. 6, pp. 117-134, 2002.
- [18] E. M. Wenzel, F. L. Wightman, and D. J. Kistler, "Localization with non-individualized virtual acoustic display cues," presented at SIGCH Conference on Human Factors in Computing Systems: Reaching through Technology, New Orleans, LA, USA, 1991.
- [19] G. D. Romigh and D. S. Brungart, "Real-virtual equivalent auditory localization with head motion," *J. Acoust. Soc. Am.*, vol. 125, pp. 2690, 2009.
- [20] D. S. Brungart, G. D. Romigh, and B. D. Simpson, "Head-related transfer function enhancement for improved vertical-polar localization," *J. Acoust. Soc. Am.*, vol. 125, pp. 2691, 2009.