

MASTER

Designing an auditory display to facilitate object acquisition in haptic virtual 3D environment

Crommentuijn, K.J.

*Award date:*  
2006

[Link to publication](#)

**Disclaimer**

This document contains a student thesis (bachelor's or master's), as authored by a student at Eindhoven University of Technology. Student theses are made available in the TU/e repository upon obtaining the required degree. The grade received is not published on the document as presented in the repository. The required complexity or quality of research of student theses may vary by program, and the required minimum study period may vary in duration.

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain

**Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

**Designing an Auditory Display  
to Facilitate Object Acquisition in  
a Haptic Virtual 3D Environment**

**Koen Crommentuijn**  
560828

1ste begeleider : dr. D. J. Hermes (HTI, TU/e)  
2de begeleider : dr. W. A. Ijsselstein (HTI, TU/e)  
extern begeleider : F. Vinberg (HCI, KTH)

**Master Thesis**

**Human-Technology Interaction  
Department of Technology Management  
Technische Universiteit Eindhoven**

November 2006

**NIET  
WILLENBAAR**

## PREFACE

To finish my Master degree in Human-Technology Interaction I was not looking for an average project. I knew a long before the search started that I wanted to go abroad. It was Kerstin Severinson Eklundh, professor and research coordinator of the Human-Computer Interaction (HCI) group at the Royal Institute of Technology in Stockholm, who made it all possible for me. She challenged me with a project that would use state-of-the-art technology to enable visually impaired persons to interact with virtual 3D environments. After seeing what was being done at the HCI group with my own eyes, the decision was easily made. For eight months I worked, lived and made friends in Stockholm, Sweden.

My project combined the interest of two researchers at the HCI group. That of Eva-Lotta Sallnäs, who focuses on how visually impaired persons can collaborate with others persons in a touchable virtual environment. And that of Fredrik Winberg, who focuses on how sound can be used to create a user interface for visually impaired persons. Fredrik Winberg became my supervisor and gave advice throughout the whole project. Eva-Lotta Sallnäs also made a large contribution; she was always available and willing to give advice.

When my time at the Royal Institute of Technology was up, I still had to write my Master Thesis. During this process I was closely guided by Dik Hermes and Wijnand IJsselstein, as my primary and secondary supervisor at the Technische Universiteit Eindhoven. Their advice helped me write a better thesis.

I've already mentioned their names, but I, explicitly, want to express my gratitude to the following people for making my Master project a success: Kerstin Severinson Eklundh, Fredrik Winberg, Eva-Lotta Sallnäs, Dik Hermes and Wijnand IJsselstein.

Finally I would like to thank my parents, Jan en Ria Crommentuyn, for their unconditional support during my education. Without them, none of this would be possible.

# **CONTENT**

**Master Thesis**

**Appendix I:**  
Procedure of the first user study

**Appendix II:**  
Short Paper

## **MASTER THESIS**

Designing an Auditory Display to Facilitate Object Acquisition in a Haptic  
Virtual 3D Environment

# Designing an Auditory Display to Facilitate Object Acquisition in a Haptic Virtual 3D Environment

**Koen Crommentuijn**

Human-Technology Interaction  
Department of Technology Management  
Technische Universiteit Eindhoven  
SE-100 44 Eindhoven, the Netherlands  
k.j.crommentuijn@tue.nl

**Fredrik Winberg**

Department of Human-Computer Interaction  
School of Computer Science and Communication  
Royal Institute of Technology  
SE-100 44 Stockholm, Sweden  
fredrikw@csc.kth.se

## ABSTRACT

Finding objects in a virtual 3D environment is difficult when only the sense of touch can be used. For visually impaired persons the addition of sound is a possible solution. To investigate how sound can facilitate object acquisition, five different auditory displays were designed. The performance of these auditory displays was evaluated in two user studies with seven visually impaired persons. The results show that information on the position of objects in relation to the cursor is necessary. Different characteristics of the auditory displays relevant to object acquisition are discussed as well as the design of an auditory display in general.

## Categories and Subject Descriptors

H.5.2 [Information Systems]: User Interfaces - *Auditory (non-speech) feedback, Haptic I/O.*

## General Terms

Design, Human Factors.

## Keywords

Auditory display, haptic, multimodal, object acquisition, virtual 3d environment, visual impairment.

## INTRODUCTION

Finding objects in space is difficult for people with a visual impairment. Without vision, the ability to identify, localize and grab for objects is limited. This problem exists both in real and in virtual environments. In virtual environments, however, it is possible to fully control the information presented, and to use abstractions and a non-realistic presentation to enhance the interaction. Sound and haptic feedback could be combined to enable visually impaired persons to successfully interact with complex and dynamic virtual 3D environments. Research on this topic is scarce, but very necessary. Furthermore, the field of application for this research is larger than enabling access to virtual environments for visually impaired persons. The audio or haptic modality can be the most suitable modality for presenting different types of information (Brewster, 2003; Snibbe et al., 2001), for example, an auditory signal representing a fire alarm, or

force feedback in a steering wheel. The auditory and haptic modality are also of interest because it can reduce visual overload (Brewster, 2003; Tang et al., 2005), for example, by reducing the number of warning lights on a airplane control panel by having a computer read aloud the warnings. Furthermore, the combination of modalities can improve the experienced realism and immersion with virtual environments (Alcañiz, Lozano & Rey, 2004).

## The Haptic Modality

The haptic modality is generally defined as the combination of the cutaneous and kinesthetic senses (Oakley et al., 2000). The cutaneous sense relates to the perception of skin sensations, like pressure, temperature and pain (Oakley et al., 2000). The kinesthetic sense relates to the perception of the position, orientation and motion of limbs, along with the associated forces (Srinivasan & Basgodan, 1997). Haptic feedback combines tactile feedback, relating to the cutaneous sense, and force and proprioceptive feedback, both relating to the kinesthetic sense (Burdea, 2000).

Unique to the haptic sense is its bi-directionality. Not only do we perceive the environment we contact, we interact with the environment. Gibson (1962 in Lederman & Klatzky, 1987) noted two different types of haptic perception; passive haptic perception, when an object moves against skin, and active haptic perception, when an object is actively touched. Especially the latter is important for object recognition (Lederman & Klatzky, 1987). The human hand is specially equipped for interacting and sensing the immediate environment, having 22 degrees of freedom, controlled by approximately 38 muscles, and incorporating an estimated 17.000 mechanoreceptors (Zecca et al., 2002). Lederman and Klatzky (1987) identified six different exploratory procedures of how we use our hands to obtain haptic information (see figure 1 on the next page). Each exploratory procedure, written in bold, is aimed at retrieving one or more properties of an object, written between parentheses.




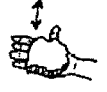


	<b>Lateral Motion</b> (texture)		<b>Pressure</b> (hardness)
	<b>Static Contact</b> (temperature)		<b>Unsupported Holding</b> (weight)
	<b>Enclosure</b> (volume, global shape)		<b>Contour Following</b> (volume, exact shape)

Figure 1: Six exploratory procedures (adapted from figure 1 in Lederman & Klatzky, 1987).

### Haptic devices

Various devices have been developed providing haptic feedback. In some cases tactile feedback was added to an already existing device, such as the FEELit mouse introduced in 1997 (Burdea, 2000). In other cases a device was specially designed to provide cutaneous information, for example, the refreshable Braille display that was introduced during the 1970's. The first system giving force feedback was a tele-operation system designed for nuclear environments in 1954 (Burdea, 2000). Computer controlled force feedback became only available during the 1990's (Burdea, 2000). Iwata (2003) differentiates four types of devices that provide kinesthetic feedback to the hands; passive props, object-orientated force displays, exoskeleton force displays, and tool-handling force displays. This study only considers tool-handling force displays, as they currently are the most practical and unrestrained devices for exploring haptic virtual 3D environments (Iwata, 2003). The first tool-handling force display that became commercially available was the PHANToM Haptic interface, developed by Massie and Salisbury in 1994 (Massie & Salisbury, 1994). Since then, various PHANToM haptic devices have been introduced, having three or six degrees of freedom for input (x, y, z, yaw, pitch, and roll) and three or six degrees of freedom in presenting force feedback. The PHANToM devices are currently one of the most popular commercially available haptic interfaces (Iwata, 2003).

### Advantages of using a tool-handling force display

Tool-handling force display devices like the PHANToM Desktop provide a novel and natural way of interacting with a virtual 3D environment. I will present a short summary of advantages of using a PHANToM or comparable haptic device to interact with a virtual 3D environment. First of all, a large amount of *realism can be added* to a virtual environment (Alcañiz, Lozano & Rey, 2004). Also the actions that the user can perform in a virtual environment become similar to performing real actions. Users can *use more of their everyday knowledge*

in virtual environments, for example, pressing a door handle down to open a door. In addition, *Immersion and the sense of presence in a virtual environment can be increased*, which can improve the performance of users and reduce the number of errors made (Sallnäs, 2004). Because touch is a very personal, socially loaded, modality, haptic feedback can *socially enrich interactions* with a computer, or with another person in a computer mediated setting like a videoconference (MacLean, 2000). Haptic feedback is also *very salient*, making it a suitable modality for sending information that needs direct attention (Subramanian et al., 2005). Additionally, the human haptic system is *well attuned to time critical actions*. Tasks that require accurate timing, for example, handing over objects to another user, can be facilitated (Sallnäs, 2004). Another advantage is that the use of a spatial input device *adds another memory system* to the interaction. Humans have the ability to remember the position of their limbs accurately and for a long period (Clark & Horch, 1986). An obvious example is the memorization of the keys of a piano or of a computer keyboard. Moreover, a complex operation, when often rehearsed, can become a mental effortless or subconscious process, for example riding a bike or driving a car.

### Limitations of using a tool-handling force display

Haptic feedback is perceived through direct contact with the immediate environment. With the exception of heat, which can travel, no haptic information is available for objects that we are not in direct contact with. Without the other senses, an environment has to be physically explored to know what is there. With the PHANToM haptic device users have to grab a stylus or place one of their fingers in a thimble, a holder for the fingertip, in order to 'touch' a virtual environment. Force feedback is applied to the tip of the stylus or to the tip of the thimble. With the PHANToM haptic device a virtual environment is explored with only one point of contact. Picking up and enclosing objects is thus not possible when using only one device. Inside a virtual environment the point of contact is very small, represented by a cursor. Normally, this cursor is very small, comparable to the tip of a ballpoint. With only a single point of contact that also is small, it is very hard to find small objects in a virtual environment and objects are easily missed (Sjöström, 2001a; 2001b).

Then, when an object is found, the haptic information perceived needs to be made sense of; the object needs to be identified. Both in vision (Goldstein, 1999) and in touch (Klatzky & Lederman, 1999), objects are identified mainly by their geometric properties. Compared to vision, the haptic system is slow and less accurate in the identification of objects (Klatzky & Lederman, 1999). Generally, the exploratory procedure of enclosure is used to quickly assess the geometric properties of an object. But with one point interaction systems, enclosure is not possible, which further limits the identification of objects.

Next, when objects are found and identified, they need to be memorized and placed in a mental representation of the virtual environment. It takes a lot of time and effort to build an accurate mental representation of a virtual environment using haptic exploration only (Wall & Brewster, 2004). And when objects are moving, it becomes increasingly difficult to form an accurate mental representation. Since users lack this continuous overview of a virtual environment, it is important that reference points are presented to facilitate navigation and the forming of a mental representation (Sjöström, 2001a).

Another limitation of the haptic modality is its limited capability of presenting absolute values. Forces are not perceived as values in Newton, but as weak or strong. Absolute values can be perceived through abstraction, through symbols. Generally, absolute values are presented by numbers, either written or spoken. Especially for visually impaired persons, Braille, a tactile version of text characters, was designed. But with a one point of contact interaction device, the dots making up a Braille character can no longer be perceived simultaneously, cancelling its advantage.

#### *Overcoming limitations*

To full exploit the advantages that haptic devices offer to visually impaired persons, the limitations have to be dealt with. To overcome some of the limitations, researchers have suggested several virtual haptic tools, for example, an enlarged cursor or attractive forces exerted by objects (Sjöström, 2001b). But virtual haptic tools can only partly compensate the limitations of unimodal haptic interaction. A different approach is to create a multimodal virtual environment, combining the haptic modality with another modality available to visually impaired persons, the auditory modality. We will investigate the auditory modality to overcome the limitation of finding objects in a haptic virtual 3D environment.

### **The Auditory Modality**

Sound is the perceptual experience of acoustic energy. The human auditory system is a remarkable system that can transform complex intertwined acoustic energy into a sense of pitch, loudness, timbre, duration and direction. It has its own way of grouping and differentiating sounds. We do not hear a series of notes when listening to a song, but a continuous piece of music. Still, we can also identify the different musical instruments and, in most cases, understand the lyrics.

Audio can convey information in four basic forms; *speech*, *non-realistic audio*, for example, an incoming message warning, and *realistic audio*, for example, the sound of a car engine in a computer game, and *music*. The earliest complete text-to-speech system for English dates back to 1968 (Klatt, 1978 in Lemmetty, 1999). During the 80's, realistic and non-realistic audio started to be

considered as an information pathway for human-computer interaction (cf. Gaver, 1986). It was also during that time, in 1984, that a loudspeaker started to be integrated into the personal computer. The soundcard made its entrance in 1988. During the early 90's the interest in auditory displays increased, exemplified by the founding of the International Community for Auditory Display (ICAD) in 1992. An auditory display presents information in auditory form and is the auditory counterpart of the visual display.

#### *Using Sound to overcome haptic limitations*

Speech is very suitable for naming objects and presenting absolute values. However, the slow rate of presentation, due to the serial nature of speech (Brewster, 1994), makes it less suitable for presenting continuous feedback on the position of objects. Therefore, non-speech audio is further investigated for facilitating object acquisition in haptic virtual 3D environments.

A general advantage of sound is that it can be heard over distance. Auditory information is perceived from sounds sources in all directions. Sounds are also rapidly detected and processed by the human brain (Kramer, 1994). Multiple sounds can be processed at once, called parallel listening (Kramer, 1994). When we hear our name being called in a conversation we are not attending, is an example of parallel listening. Some sounds can be attention grabbing, for example, the horn of a car, yet other sounds can blend into the background, such as the sound of the air-conditioner (Kramer, 1994). But when a sound that blended into the background suddenly changes, this change is often noticed.

Another advantage of non-speech sound is the wide variety of sound characteristics, such as pitch and roughness, that humans can discriminate. When a sequence of sounds is played, we perceive more than individual sounds. Sounds are grouped and discriminated according to auditory Gestalt principles like proximity and similarity. When hearing a sequence of sounds, we perceive characteristics as tempo, rhythm and melody. These perceptual dimensions can be used to present information.

#### *Design issues with non-speech Audio*

Using sound also has its limitations. Kramer (1994) discussed various difficulties concerning auditory displays. The main problems Kramer (1994) discussed are: (1) *the low resolution of many auditory variables*. Many dimensions of sound have a low resolution. For example, only a few values can be unambiguously presented with volume (Buxton, Gaver & Bly 1991 in Brewster, 2003). (2) *The lack of absolute values*. Without using speech, it is hard to represent absolute values. It is extremely hard to judge the absolute value of a dimension of sound. It is much easier to perceive changes of an



auditory dimension. (3) *The lack of orthogonality*. Many perceptual dimensions can be discriminated in sound, but these dimensions are seldom independent. A change in one dimension will affect the perception of the other dimension. (4) *The absence of persistence*. Acoustic energy travels and dissipates over space. A sound therefore only exists for its duration in one point of space. Users must rely on their memory to understand and compare auditory information. (5) *The annoyance users can experience from an auditory display*. Not all sounds and sequences are perceived as pleasant. Sound can be annoying and cause discomfort, for example, the sound of nails scratching across a blackboard. And while we are able to close our eyes for something we do not want to see, we cannot close our ears for a sound we do not want to hear. Another limitation, not listed by Kramer (1994), is *the mixing of different sounds*. Sound is perceived through two channels, our left and right ear. Acoustic energy originating from different sources merges together. Only through complex processes, taking place in the cochlea and our brain, we can separate multiple sounds. This processing is however limited, both physical and perceptual. For example, intense acoustical energy can mask less intense acoustical energy and when too many sound sources are present we can no longer keep them apart in our mind. Finally, Winberg (2001) reports *the lack of conventional use of sound*. As of yet there is little convention on the mapping of auditory dimensions to information. Users cannot draw upon previous experience when learning an auditory display, unless natural sound effects are used.

#### *Previous auditory displays*

Researchers have designed various auditory displays that give visually impaired persons access to a graphical user interface. Winberg (2001) discriminated three main models on which auditory interfaces are based; the linear, hierarchical and spatial model. With the linear model, auditory information is presented sequentially. Screen readers are examples of auditory interfaces that are based on the linear model. With a hierarchical model, users step through a tree-structured list of which the items are read aloud. Examples of a hierarchical model can be found in the Mercator project (Edwards, Mynatt & Stockton, 1994) and in the auditory interface designed by Savidis et al. (1996). With the spatial model, objects are placed in a virtual area similar to a graphical user interface. The spatial model is the only model that represents a virtual environment in its natural form. While the linear and hierarchical models ensure that all information is easily accessible, the spatial model is the only model that allows a natural interaction with a virtual 3D environment.

One of the first systems that added an auditory user interface to the graphical user interface was the SonicFinder by Gaver (1989). With the SonicFinder, auditory icons were mapped to system events, such as

selecting a file and copying a file. Auditory icons are everyday sounds that metaphorically represent actions and events, for example, the sound of a glass being filling with water used to represent a progress bar. The SonicFinder was only designed to augment the graphical user interface. The SonicFinder did not address the navigational issues when visually impaired persons try to use the interface.

Soundtrack, created by Edwards (1989), was one of the first spatially based auditory user interfaces specially designed for visually impaired persons. Soundtrack was a text editor that was fully accessible by auditory feedback. Edwards (1989) combined non-speech sound for navigating the user interface with speech synthesis for menu's and written text. Finding objects on the user interface was simple, because the workspace of the interface was divided into eight blocks and each whole block represented an object. The use of the non-speech sound did not work as expected. Users counted the sounds that were played when entering a new object, instead of inferring which object the sound indicated (Pitt & Edwards, 1991).

To create a more general application, Pitt & Edwards (1991) investigated how objects could be located by sound with the use of a mouse. First, a visual navigation experiment was conducted, investigating how well users could use distance-, directional- and combined information to find an object. The results show that objects are approached faster using directional information than using distance information. They also found that when directional information and distance information was presented simultaneously but separately, the localization of objects took much longer than when the directional and distance information were presented in their natural combined form, when presenting a cursor and target object. Pitt & Edwards (1991) then replaced the visual cues with auditory cues. An object was represented by a synthesized cello tone. Volume was used to represent distance and was used independently or combined with stereo panning, which represented the horizontal direction, "the intended analogy was that of the cursor acting as a microphone" (Pitt & Edwards, 1991). Pitt and Edwards (1991) also experimented with a non-linear volume increase, having the volume increase faster when the cursor comes closer to the target. This was done because, in a pilot, users were fast in moving the cursor close to an object but took a longer time to finally acquire the object. Their results show that both the use of a non-linear increase and the use of stereo panning, reduced the time necessary to acquire an object.

Mereu and Kazman (1996) experimented with audio enhanced 3D interfaces. The x, y, and z-coordinates of the cursor and of the target were mapped, in the same order, to panning, pitch and volume. Standard, the user would

hear a sound which represented the position of the cursor, but when pressing the left mouse button he or she could hear a sound which represented the position of the target. Mereu and Kazman (1996) found that users could accurately acquire a target by comparing sounds.

Magnusson and Rasmus-Gröhn (2005) investigated two different auditory user-interfaces integrated in a haptic virtual 3D environment. First, they investigated a 3D version of the virtual microphone method from Pitt and Edwards (1991), mapping virtual sounds sources to the position of objects and using the cursor as a microphone to listen to the virtual environment. And second, they investigated the use of virtual “sound beams” extending from objects along the z-axis. The virtual microphone method was reported as intuitive to use, but nothing was reported on the use of virtual sound beams (Magnusson & Rasmus-Gröhn, 2005).

### This Study

Only a small amount of projects have combined haptic and auditory feedback for exploring non-visual virtual 3D environments. Haptic devices, like the PHANTOM Desktop, offer a natural and intuitive way of interacting with a virtual 3D environment. To enable visually impaired users to take advantage of this new interaction device, we need to overcome the limitations of unimodal haptic feedback. The addition of sound opens up many possibilities, but only a very limited amount of research is available on how sound can be used to overcome these problems. Therefore a qualitative investigation, on how sound can facilitate the acquisition of objects, was carried out.

Five different auditory displays were implemented in a virtual haptic environment developed for the purpose of this investigation. Each of the five different auditory displays used different combinations of auditory dimensions to represent spatial information. An auditory display maps information onto one or more perceptual dimensions of sound. Creating a complete and consensual classification is almost impossible due to the great number and ambiguousness of the perceptual dimensions. Though, a reasonable consensus exists for several basic dimensions. Most books on sound perception differentiate *pitch*, *loudness* and *timbre* (Goldstein, 1999; Sekuler & Blake, 2002). While pitch and loudness are two well defined concepts, timbre seems to be a collection of everything not defined by pitch and loudness (Brewster, 1994), including, for example, roughness and brightness. Two generally used additions to these three dimensions are *duration* and *location*. Also, when sounds are played over a period of time, various other dimensions can be discriminated. This study considered the *start*, *tempo*, *dynamics*, and *composition* of sounds (figure 2); the start of a sound refers to its origin in time, tempo refers to the time between sounds, dynamics refer to the perceived

changes between sounds over time, for example, a rising pitch or decreasing loudness, and composition refers to the combination of different sounds.

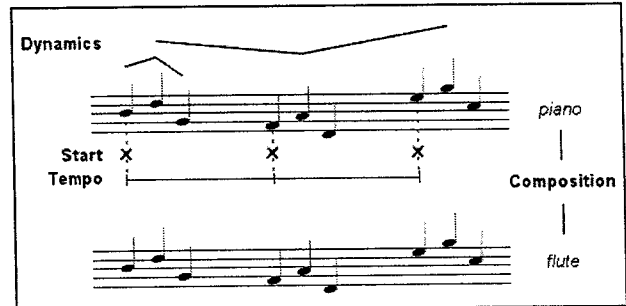


Figure 2: Dimensions in a sound sequence.

### The Virtual Environment

The haptic 3D virtual environment designed for this investigation was bounded by a box measuring 100 x 100 x 80 millimeters (see figure 3). The objects in the virtual environment were cubes measuring 8 x 8 x 8 millimeters.

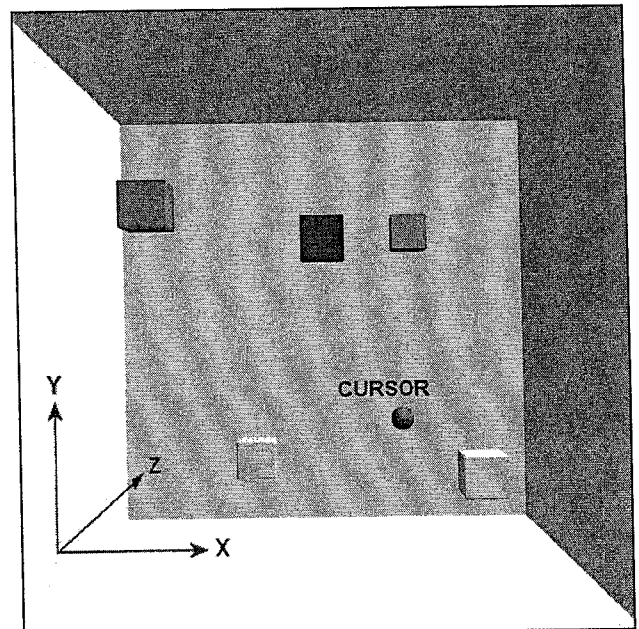


Figure 3: Visual representation of the virtual environment.

The cubes could be moved by making contact with the cursor and subsequently pressing the button on top of the stylus. The cubes could not rotate and no gravity was simulated in the virtual environment. The cubes stayed in the exact position where they were created by the system or placed by the user. The initial position of a cube was randomly generated by the system.

#### ***AD1: Coordinate-based Auditory Display***

The first of the five auditory displays was the only auditory display that sonified the absolute position of objects. The auditory presentation only changed when objects were moved and not when the cursor was moved. A virtual auditory scanner, represented by a low pitched buzzing sound, continuously swept from left to right through the virtual environment in approximately 2,5 seconds. When the scanner passed an object, a bell sound was played to represent that object. Characteristics of the bell sound were transformed to represent the position of the object. Five panning values (left, left/center, center, right/center, and right) were mapped to the x-coordinate. Pitch was mapped to the y-coordinate and a cross-fade from an original to a high-pass filtered sample was mapped to the z-coordinate. The high-pass filter made the bell sample sound thin and subdued. Volume, as used in the experiment of Mereu and Kazman (1996), was intentionally not chosen to represent the z-coordinate, because it interfered with panning.

#### ***AD2: Nearest Path Auditory Display***

The second auditory display sonified the direction and distance from the cursor to the nearest object. Again a bell sound was used to represent the objects. When a different cube became the closest cube to the cursor, a short double bell sound was played to indicate that a new path was being sonified. Discrete transformations were used to represent the path from the cursor to the object. Panning (left, right and center) was used to indicate if the cursor should be moved to the left, to the right or kept in the same horizontal position. Pitch (high, low, and medium) was used to indicate if the cursor should be moved upward, downward or kept on the same height. And a high-pass filter (on or off) was used to indicate whether the cursor should be moved backward or not. The bell sounds were played repetitively. The time between repetitions decreased logarithmically as the distance from the cursor to the object decreased. This created an auditory space with a low resolution far from the objects and a high resolution close to the objects, as suggested by the investigation of Pitt and Edwards (1991).

#### ***AD3: Virtual Microphone Auditory Display***

In the third auditory display the haptic cursor acted like a virtual microphone through which the user can listen to the virtual room, similar to an auditory display investigated by Magnusson and Rasmussen-Gröhn (2005). The objects, represented by different drum loops, were positioned in 3D auditory space corresponding to their position in the virtual environment. This was done by applying Head Related Transfer Functions (HRTFs) in real-time to the drum loops. A HRTF imposes an interaural time difference and integrates absorption, reflection and diffraction effects caused by the head, torso and pinnae of a general user. The HRTFs, used in this investigation, were programmed into the hardware of the

soundcard. The 3D auditory space was set-up in such a way that objects close-by were heard loudly, objects at a small distance were still audible and objects far away could not be heard. The drum loops used were selected in advance based on the quality of synthesis, broadness of spectrum and distinctiveness of rhythm.

#### ***AD4: Prioritized Notes Auditory Display***

The fourth auditory display was based on the idea that objects close to the haptic cursor are of larger interest than distant objects. This auditory display represented each object with a different piano note. Together these notes formed a pleasant sounding chord. Distant objects were sonified simultaneously once or twice every 1.8 seconds depending on their distance. Objects close to the cursor were sonified independently of the tempo of the far objects. In this manner, objects far from the cursor blended together, and objects close to the cursor stood out. For objects close-by, the time between repetitions decreased logarithmically as the distance from the cursor to the object decreased identically to how distance was represented by the nearest path auditory display. Similarly to the virtual microphone auditory display, all piano notes were positioned in 3D auditory space. The 3D auditory space was however modified, in such a way that all notes played were audible, even when the distance between the cursor and an object was at its maximum.

#### ***AD5: Pointing-based Auditory Display***

The fifth auditory display transformed the stylus of the PHANTOM Desktop into a pointing device. The user had to point the stylus through the virtual environment and when the stylus was pointing at an object, a bell sound was played repetitively. The bell sound was transformed, similarly to the coordinate-based auditory display, to represent the absolute position of the object. The time between repetitions, similarly to the nearest path auditory display and the prioritized notes auditory display, decreased logarithmically as the distance from the cursor to the object decreased. When not pointing directly at but close to an object, a sound was played that indicated how users should rotate the pen to point directly at the object. This guiding sound can be described as an 'ahhh' sound. Three pitch values (high, middle, and low) were mapped to rotation around the x-axis, also called pitch rotation. And three panning values (left, center, and right) were mapped to rotation around the y-axis, also called yaw rotation. The guiding sound was designed to make it easier to find objects by giving auditory feedback on a larger area than the cube itself.

Table 1: The mapping of information to auditory dimensions.

	AD1 (absolute)	AD2 (cursor related)	AD3 (cursor related)	AD4 (cursor related)	AD5 (cursor related)
<b>Sound</b>					
pitch	y-value	y-direction	-	identification	y-value / pitch rotation
loudness	-	-	-	-	-
timbre	z-value	z-direction	-	-	z-value
duration	-	-	-	-	-
<b>location</b>					
lateral	x-value	x-direction	-	-	x-value / yaw rotation
3D surround	-	-	direction / distance	direction	-
<b>Sequence</b>					
start of a sound	x-value / identification	-	-	-	-
tempo	-	distance	identification	distance	distance
dynamics	-	-	identification	-	-
composition	-	-	identification	-	-

**The five auditory displays compared**

Each of the five auditory displays relied on different combinations of auditory dimensions to represent spatial information. The mapping of information to auditory dimensions is presented for each auditory display in table 1. This table also shows if, and by which auditory dimensions, objects could be identified and discriminated from each other. Since each auditory display was based on a different principle and relied on different auditory dimensions, it was expected that some auditory displays would be more intuitive to use than others. However, all of them were expected to be learned within a relative short amount of time. When learned, it was expected that objects were detected quickly with auditory displays that presented multiple objects at once; AD1, AD3 and AD4. With AD2 and AD5, where only one cube could be heard at a time, it was expected that it would be easy to move towards a detected object, but detecting them would be more difficult.

**The haptic magnetism tool**

Developers of haptic applications generally resorted to haptic tools or widgets to aid target acquisition. Crosshairs, magnets, and haptic cursor enlargements are examples of suggested haptic tools (Sjöström, 2001b). It is still unknown how auditory displays compare to haptic tools, when both are designed for facilitating target acquisition. It is also still unknown whether an auditory display can be usefully combined with a haptic tool. To get a general answer to those questions, a form of haptic magnetism was implemented along with the auditory displays. With haptic magnetism enabled, objects exerted small attractive forces upon the haptic cursor, pulling the hand of the user towards the objects. The added forces were intentionally kept small, to avoid interference with the movements of the user.

**METHOD**

**Design**

This research investigates how auditory feedback can support the detection and acquisition of objects in a haptic virtual 3D environment. The five different auditory displays and one haptic tool, implemented in a haptic virtual environment, were evaluated in two qualitative user studies. In the first user study, participants used all five auditory displays and the haptic tool. In the second user study, participants used only two auditory displays and the haptic tool.

Both user studies had, for each auditory display, an introductory phase, a question phase, a task phase and an interview phase. During the introductory phase, participants got instructions and tried out a randomly selected auditory display. During the question phase, participants used this auditory display to answer questions about the global lay-out of the objects within the virtual environment. During the task phase, participants were asked to move cubes to various places within the virtual environment. During the interview phase, participants were asked to comment on the auditory display they just used. In the first user study participants were additionally asked to rank the auditory displays on usefulness and pleasantness. When there still was time available after exploring and, in the first user study, ranking the auditory displays, the haptic tool was evaluated.

The whole procedure was captured on video. The following information was extracted through video analysis: (1) the percentages of incompleteness and the time to complete questions and tasks, (2) remarkable events, abilities and disabilities of the participants, and (3)

relevant remarks articulated by the participants during the evaluations and the interview phase. The rankings on usefulness and pleasantness were immediately written down when given by the participants during the first user study.

We will use the quantitative data to indicate how well the different auditory displays performed. The qualitative data will be used to identify the characteristics of each auditory display and the haptic magnetism tool that are responsible for its performance.

## Participants

One female and two male participants, ages 30, 54, and 55, took part in the first user study. All three participants were born with a visual impairment, and have been legally blind for a minimum of 10 years. One of the three participants reported some previous experience with a haptic device similar to the PHANToM Desktop. Four additional participants, two male, and two female took part in the second user study. Except one female, who still had low vision (10/100 unaided), all four participants were legally blind from birth. Their age varied between 26 and 59. One of the four participants reported some previous experience with a haptic device similar to the PHANToM Desktop.

## Apparatus

The virtual environment was run on a modern desktop computer with Microsoft Windows XP Professional installed as operating system. A Sound Blaster X-Fi XtremeMusic from Creative was installed and used to play audio. Its Creative MultiSpeaker Surround 3D feature (CMSS-3D) was used to produce 3D Surround Sound over a set of Sennheiser headphones. The PHANToM Desktop haptic device of SensAble Technologies was used as a 3D input device that simultaneously gives kinesthetic feedback.



Figure 4: Video-frame of a mixed audio/video-file.

A digital video camera and webcam were used to capture the upper body of the participant, his or her use of the haptic device, a visual representation of the virtual environment, the audio generated by the computer and the conversation between the participant and the observer. The video and audio streams were mixed together in a single audio/video file (see figure 4).

## Procedure

### *User Study 1:*

#### *Haptic Introduction*

To get familiar with haptic interaction, participants first explored, for approximately two minutes, a virtual haptic environment containing a large pyramid shape and a large doughnut shape. Participants were asked to find and identify these two shapes without the help of an auditory display or the haptic magnetism tool.

#### *Auditory Display Introduction*

Next the participant got, in random order, an introduction to one of the five auditory displays. After the introduction, participants used the auditory display in a virtual environment containing two cubes. During their exploration, participants received additional instructions on how to use the auditory display. Participants were allowed to ask questions about the auditory display at all times. After participants could successfully locate and manipulate the two cubes in the virtual environment, the evaluation was continued.

#### *Question phase*

After the introduction on an auditory display, an environment was started with a randomly three to eight cubes present. Participants were asked to describe the global position of the cubes, or how many cubes were present in the virtual environment. Two of these questions were scheduled per auditory display.

#### *Task phase*

After the question phase, a new environment was started with five cubes present. The participants were asked to move the cubes to a random wall of the room, for example left, or to move half of the cubes to a random wall and the other half to the opposite wall, for example front and back, or to build a tower of the cubes in the middle of the floor. When the participants could not find the next cube within a reasonable time, the task was ended by the observer. Three tasks were scheduled per auditory display, but generally only one or two tasks were completed due to time constraints.

#### *Interview phase*

After completing the task phase, participants were asked several questions about their experience with the auditory display they just used, for example, "Can you tell me why the tasks were easy or difficult?".

### **Ranking phase**

The process of auditory display introduction, question phase, task phase and interview phase, was repeated until all five auditory displays had been evaluated. A ten minutes break was held between evaluating the third and fourth auditory display. After evaluating all five auditory displays, participants ranked the auditory displays on usefulness and pleasantness.

### **Haptic tool**

When there still was time remaining after testing all five auditory displays, participants evaluated the haptic magnetism tool. Haptic magnetism was enabled independently or combined with one of the auditory displays. The participants were given questions and tasks identical to the ones asked during the evaluations of the auditory displays.

The total procedure was completed within approximately two hours.

### **User Study 2:**

The start of the second user study was identical to the first evaluation. But instead of randomly testing five auditory displays, two out of four auditory displays were pre-selected. The coordinate-based auditory display was excluded from the second user study, because participants could hardly use this auditory display to come into contact with a cube. In the second user study, the question phase was swapped with the task phase to see if experience from the task phase increased performance during the question phase. Six tasks and two questions were scheduled per auditory display in the second user study. Participants were given a break of ten minutes between testing the first and second auditory display. When time remained, participants also evaluated the haptic magnetism tool independently and combined with the two previously evaluated auditory displays.

The total procedure was again completed within approximately two hours.

## **RESULTS**

A total of 33 questions on global lay-out were answered during the first user study; 26 while evaluating an auditory display and 7 while evaluating haptic magnetism. In addition, a total of 26 tasks were completed; 18 while evaluating an auditory display and 8 while evaluating haptic magnetism. During the second user study, a total of 15 questions were answered, all while evaluating an auditory display. In addition, a total of 55 tasks were completed; 46 while evaluating an auditory display and 9 while evaluating haptic magnetism.

This section reports about two types of results, quantitative and qualitative. The quantitative results

indicate how well the different auditory displays aided participants in answering questions and completing tasks. It does not **prove** that one auditory display is 'better' than the other. For haptic magnetism, only qualitative results are reported, because too few trials were completed for a quantitative comparison.

## **Quantitative results**

### **Percentage of incompleteness and time taken per object**

The percentage of incompleteness was determined by the relative number of cubes that the user did not count, locate, or move. For questions, the time per object was determined by dividing the time it took to complete a question by the number of cubes present. For tasks, the time it took to locate and move a single object was directly measured during the video analysis. For this reason, more observations are available on time per object than on percentage of incompleteness.

Figure 5 (see next page) summarizes the quantitative measurements for both user studies. The figure is split into four graphs, one for the percentages of incompleteness of the first user study, one for the percentages of incompleteness of the second user study, one for the time taken per object during the first user study, and one for the time taken per object during the second user study. Each graph is divided in a question section and a task section. Per auditory display, a bar is displaying the percentage of incompleteness or the time taken per object, is given. The absolute percentages and times are displayed above each bar. The standard deviations are displayed in parenthesis next to the absolute percentages and times. The numbers of observations are given under each bar.

In the graphs of the second user study, the first auditory display is missing, because the coordinate-based auditory display (AD1) had been excluded from the second user study.

### **Answered Questions**

During the first user study, the participants answered the questions quickest and most accurate while using the coordinate-based auditory display (AD1). With this auditory display the participants did not have to explore the environment with the haptic device answer the questions. Participants also did well, though slower, using the virtual microphone auditory display (AD3). Participants had a little more difficulty answering the questions using the prioritized notes (AD4) and pointing-based auditory display (AD5). The least useful auditory display for answering the questions was the nearest path auditory display (AD2); the average percentage of incompleteness was 83% (figure 5).

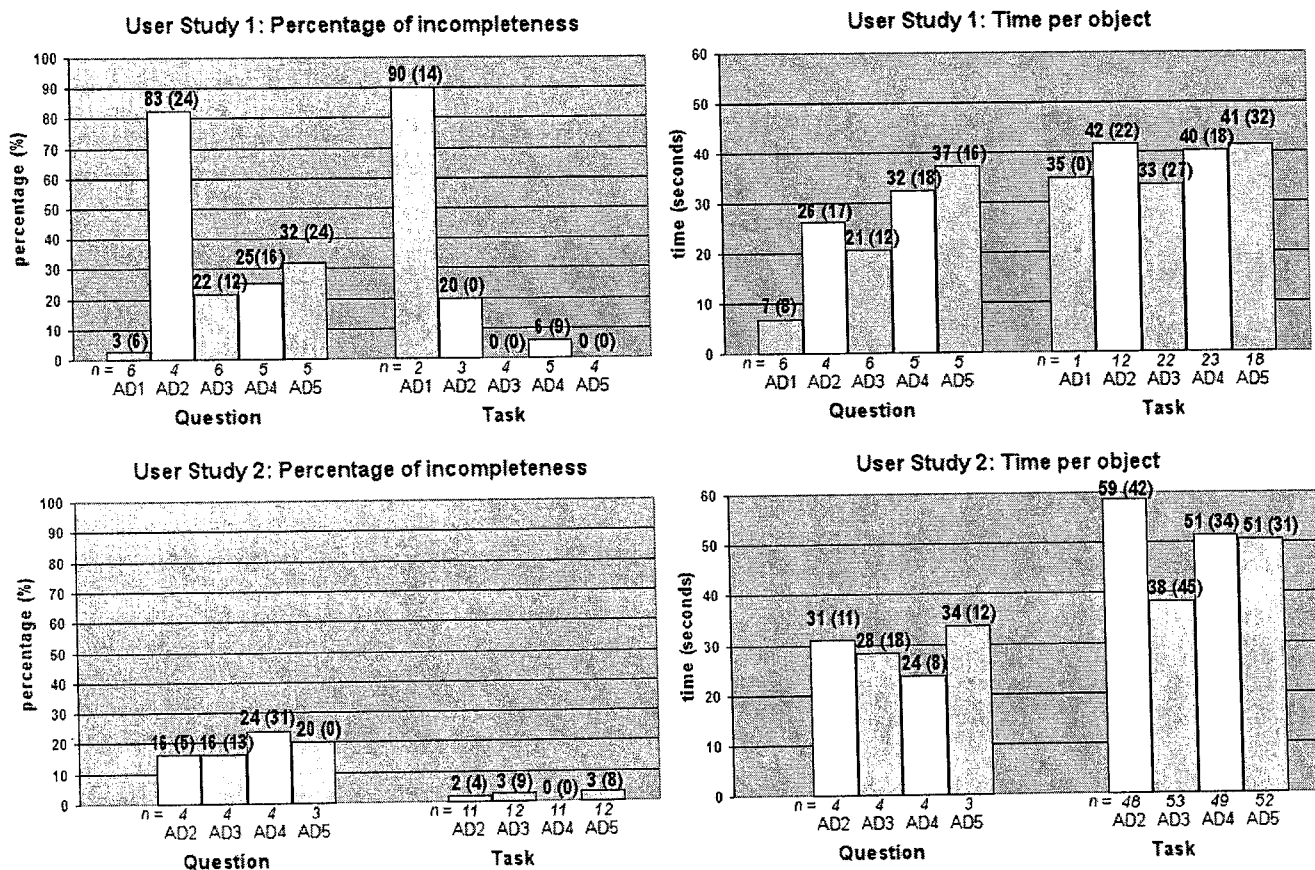


Figure 5: Percentage of incompleteness and time per object values.

During the second study, the participants had more experience using the auditory displays before answering the questions. The differences in percentage of incompleteness and time per object became very small between the auditory displays. Questions were answered accurately, regardless of which auditory display was being used, even with the nearest path auditory display (AD2).

#### Completed Tasks

The participants were unable to successfully complete the tasks while using the coordinate-based auditory display (AD1). Only one cube was successfully found and moved, illustrated by the 90% incompleteness bar in figure 5. For this reason the coordinate-based auditory display was excluded from the second user study. With the four other auditory displays, almost all tasks were successfully completed. No significant differences in the percentages of incompleteness between these auditory displays were found.

During both user studies, the participants located and moved the cubes fastest while using the virtual microphone auditory display (AD3). Second, in time per object, came the prioritized notes auditory display (AD4) and the pointing-based auditory display (AD5). Their

average performance was similar to each other, despite being based on very different principles. Out of the dynamic auditory displays, hereby excluding the coordinate-based auditory display (AD1), the participants took the longest time to locate and move the cubes using the nearest path auditory display (AD2).

The average time per object was higher during the second user study compared to the first user study. During the second user study, participants had to complete the relatively difficult task of building a tower from the cubes fourteen times, compared to two times during the first user study. A relatively large increase in time per object for the nearest path auditory display (AD2) was noticed between the user studies. The time per object increased from 42 seconds in the first user study to 59 seconds in the second user study. This increase can be explained by the extra time given to the participants in the second user study, when having trouble finding the next cube. Logically, this also led to a decrease in the percentage of incompleteness.

#### Perceived usefulness and pleasantness

During the first user study, the three participants ranked all five auditory displays on usefulness and on pleasantness. Rank 1 was reserved for the most, and rank

5 was reserved for the least useful or least pleasant auditory display. The average rankings for each auditory display are given in table 2.

Table 2: Avg. rankings on usefulness and pleasantness.

	Perceived Usefulness	Perceived Pleasantness
AD1: Coordinate-based	3,8	4,2
AD2: Nearest Path	3,8	3,5
AD3: Virtual Microphone	1,7	1,0
AD4: Prioritized Notes	3,7	3,3
AD5: Pointing-based	2,0	2,3

The virtual microphone and pointing-based auditory display were ranked as the two most useful auditory displays. While the use of the pointing-based and prioritized notes auditory display resulted in similar percentages of completion and times per object, the pointing-based auditory display was clearly perceived as more useful. When rating the auditory displays on pleasantness, one person remarked that it is more important that sounds are informative. Nonetheless, participants unanimously ranked the virtual microphone auditory display as the most pleasant auditory display to listen to. Most appealing to the participants was the sound of the drum loops and the naturalness of the auditory display. The pointing-based auditory display was ranked second on pleasantness. One of the reasons mentioned by a participant, on why this auditory display sounded pleasant, was that the auditory display could be made silent by pointing the stylus away from the cubes. While the coordinate-based auditory display was ranked as the least pleasant sounding auditory display, none of the three participants reported it as annoying.

## Qualitative results

### *AD1: Coordinate-based Auditory Display*

Counting the number of cubes in the virtual environment was easy for the participants. They could quickly identify the start and end of a sequence, by the low buzzing sound being panned from left to right, and by the recurring rhythm of the bell sounds. When two or more cubes had similar horizontal positions (x-values), participants were, generally speaking, still able to distinguish the different bell sounds. When asking to describe the global position of the objects, participants always started by describing the horizontal positions of the cubes. The temporal aspect of the virtual auditory scanner made it easy to estimate the horizontal position, as shown in the following participant remark; "If you only want to know if it is to the left, right or in the middle then I like the scanning method, because it moved from left to right. It was harder to listen to the pitch and the bells to find out if it was back or forward or high or low". Participants indeed had more trouble estimating the vertical position of the cubes. It was easier to relate the vertical position of one cube to another cube than to relate the vertical position of one cube to the

room. Estimating the front-back position of a cube was the most difficult for the participants. In several cases no estimation at all was given for the front-back position; "It is hard to say if they are close to the front or to the back wall", was one of the remarks. In general, participants could only distinguish depth between cubes being to the far back of the room and cubes being to the very front of the room. During the user studies, it was also noticed that the high-pass filtered sample interfered with the pitch transformation of the bell sound. This led to several wrong estimations on the vertical position of cubes.

The participants were unable to complete a single task using the coordinate-based auditory display within a reasonable time. Participants did have a general idea of where cubes were located, but the cubes could not be found with the haptic device. Participants were moving the cursor in large scanning motions, or along the walls, in the hope of bumping into a cube. During the haptic exploration, participants were relying more on luck than on the information from the auditory display.

### *AD2: Nearest Path Auditory Display*

It took participants quite a while to understand, and correctly use, the nearest path auditory display. The sound played when a different cube became the closest cube to the cursor was often misunderstood. For example, some participants mistakenly thought that the cursor was very near to an object when hearing this sound. Participants asked a variety of questions about this sound and, for three of the five participants who used this auditory display, its meaning had to be explained several times. It was also observed that participants had trouble distinguishing the different cubes; some cubes were counted twice and some cubes skipped, because participants mistakenly thought they had already counted them. During the second user study, participants did a lot better in answering the questions. They had learned how to find the cubes with the auditory display, and went to touch all cubes in order to count them and describe their position.

Both in the first and in the second user study, tasks took the longest time to complete with the nearest path auditory display compared to the other three dynamic auditory displays. Nonetheless, participants were able to successfully accomplish all tasks when given enough time. Participants were quick in using tempo as an indicator of distance. It took longer to learn and use the discrete directional cues. To four of the five participants who used this auditory display, the meaning of the directional cues had to be explained not only during the introduction, but also once or twice during the evaluation of the auditory display. Two participants reported that they wanted a more distinct difference between the levels of the directional cues.



### ***AD3: Virtual Microphone Auditory Display***

Participants needed little explanation to start working with the virtual microphone auditory display. Participants could answer the questions using the virtual microphone auditory display well, though they needed time to explore and find the different drum loops in the virtual environment. Occasionally a cube was missed. Some participants did not always touch the cubes to count them or describe their location. This did almost never happen with the other auditory displays or with the haptic magnetism tool. In approximately six cases, when a cube was near to a wall, the wall was mistaken for the cube.

Participants reported that they could easily distinguish most of the different drum loops. But when having to compare drum loops from memory it was more difficult. One participant, for instance, had to start over counting, because he forgot which sounds he had already heard. Some participants learned the drum loops after a few tasks, and reported during a task to be searching for a specific drum loop, for example, "I knew what sound I was looking for, for the fifth one". Other participants, however, reported they wanted the drum loops to sound more differently. It was also noticed that some drum loops could be heard over greater distances, which annoyed one of the participants. She reported that, because of this, it was harder to find drum loops that did not carry as far.

With the virtual microphone auditory display the cubes were quite easily discovered. But from discovering to coming in contact with the cube took longer. This was expressed by one of the participants: "I hear it all right, but to get it with touch is more difficult". Participants noticed when they got nearer to a cube, but the direction to move in was more difficult. Hearing if a cube was to the left, middle or right of the cursor was relatively easy with 3D Surround Sound. Hearing if a cube was to the front-back of the cursor, or above-below the cursor was difficult. Nevertheless, one participant reported he could hear and use these cues from the 3D Surround Sound.

### ***AD4: Prioritized Notes Auditory Display***

It took some time for the participants to get a sense for the tempos and the blending of the notes, but during the reflection phase the prioritized notes auditory display was reported as being easy and intuitive to use. Participants were not able to distinguish all individual notes when played in a chord. This made it necessary to explore the virtual environment in order to identify all the different cubes. Notes were easy to compare directly, but when having to compare them from memory, it became more difficult. Participants were occasionally unsure if they had already come across the cube they just found. A note played in an asynchronous faster tempo was easily distinguished from the other notes. When participants noticed a cube, they were relying more on the distance information, than on the directional information provided

by the 3D Surround Sound. Participants were observed moving the cursor back and forth in multiple directions, listening to the change in tempo and then finding the right direction to move closer to a cube.

### ***AD5: Pointing-based Auditory Display***

Generally, participants needed some time and extra explanation to understand and use the pointing-based auditory display, but afterwards several reported this auditory display as being easy to use. Participants had to actively scan the virtual environment room to find the cubes. No major strategy could be discovered in the scanning motions of the participants. One person remarked: "Some times you find them and sometimes you don't. I don't know if there is any reason for it". A few participants tried to scan the virtual environment using a probing motion, or forward longitudinal motion, which is useless for discovering new objects. One participant rotated the stylus sideward when scanning the virtual environment and was able to successfully locate several cubes. Without a structural strategy participants were still able to scan a great deal of the virtual environment and find most cubes. Some places in the front of the virtual environment were harder to point at, because the stylus needed to be pulled to the very front of the workspace or rotated into an unusual angle. Participants also had trouble differentiating cubes that were positioned close together or in one line.

The guiding tone, when pointing near a cube, was not used in the same manner by all participants. Most of the participants began by using the guiding tone as a border around a cube. As time progressed, two of the five participants started to use the directional cues given by the guiding tone. This was demonstrated in the way participants closed in on a cube when they located one. Since moving the stylus in a straight line towards a cube was not easy, participants occasionally had to re-orientate the stylus to point directly at a cube again. While three participants pulled back on the stylus, two started to use the directional cues to rotate the pen, to point directly at the cube again.

### ***Haptic Magnetism***

In total, 24 trials (7 questions, 17 tasks) were carried out with haptic magnetism enabled; 8 trials with haptic magnetism independently and 16 trials with haptic magnetism combined with an auditory display. In general, participants were positive about the haptic magnetism. One of the participants remarked for example, "I liked it, it was easy". Participants were particularly positive when haptic magnetism was combined with one of the auditory displays. Several positive remarks were made about using an auditory display to get near to a cube and then letting the haptic magnetism pull you in. For example, "When you hear you are not so far you let the pen show you were the object is" and "I heard the sound getting stronger so I

knew what direction to move, and then when I got close enough I got sucked in". Three tasks were completed remarkably fast in which haptic magnetism was combined with the virtual microphone auditory display, for example, finding and moving five cubes to the left wall of the virtual environment in 30 seconds.

There were also problems when using the haptic magnetism tool. Without an auditory display, cubes can only be distinguished by their position. Participants, therefore, could not always remember if they had touched a particular cube already, especially when several cubes were located close together. Combining haptic magnetism with an auditory display did not always speed up a task either, for example, only being able to move three out of four cubes to the bottom of the virtual environment in a time of 215 seconds. At times haptic magnetism interfered with the desired movements of the participants. Participants remarked: "In this case I would probably appreciate not to have the magnetics interfere with my search" and "the magnetic force perhaps wants me in another direction than I want". With haptic magnetism enabled, some participants also started to make more errors while trying to grab a cube in order to move it. One participant, for instance, failed to grab a cube, six consecutive times.

#### *The haptic device*

For all participants, the haptic device created a realistic haptic experience. Two of them even reported that they felt compelled to use their free hand to reach in and touch the virtual objects. The way of using the haptic device differed among the participants. Some were using the haptic device with confidence, making large and forceful movements, and others were using the haptic device in a more cautious way, making small movements and keeping the cursor in contact with a wall for continuous haptic feedback.

The stylus of the haptic device was not always grabbed in the intended way. Holding the stylus similar to holding a regular ballpoint is not self-evident. Four out of seven participants grabbed the stylus as one normally grabs a stick or a door handle. These four participants were instructed to hold the stylus in the intended way, but two of them, one immediately and one later on, switched back to their initial gripping style. The different grip did not seem to interfere much with the exploration. However, when moving the stylus to the top-back of the virtual environment, the haptic device produced a mechanical force which led to some confusion.

During the evaluations one participant experienced physical discomfort in the upper arm from using the haptic device. After half an hour this participant reported that using the haptic device was "tiring" and added: "I would suggest a support for your arm". After the

scheduled break in the middle of the user study the participant started supporting the right arm with her left arm. Soon after the break she reported: "my arm is hurting a little". The position of the haptic device was changed in a way that the left hand could be used to explore the virtual environment. It was observed that using the left arm also caused strain. The participant did not seem to perform worse after switching arms, though in general, this participant had more difficulty answering questions and completing tasks.

## **DISCUSSION**

All results indicate that the virtual microphone auditory display was the best auditory display to be used in the, here explored, virtual environment. The principle of a virtual microphone is intuitive and is easy to implement in other virtual environments. The pointing-based auditory display also performed well. With this auditory display it was harder to detect objects, but, as expected, it was easy to move in and touch the objects. A pointing-based auditory display could be very useful in virtual environments with larger objects, because they are easier to find. The prioritized notes auditory display performed equal to the pointing-based auditory display. The addition of tempo as an indicator of distance should have made this auditory display outperform the virtual microphone auditory display, which it clearly did not. This method needs further development, before it has advantages over the virtual microphone method. The nearest path auditory display also needs further development before it becomes a useful auditory display. The coordinate-based display is not suitable as an auditory display that facilitated object acquisition.

### **Designing an auditory display**

All sounds in an auditory display should have a purpose. Too many sounds will clutter the auditory space and, therefore, sounds should be presented parsimoniously. The information conveyed in an auditory display should also be easily understood by the user-group. This was, for example, not the case with the nearest path auditory display. A sound that is informative to the user is also less likely to be perceived as annoying. The use of rich natural sounds has the potential to further increase the appeal of an auditory display. Also important to the pleasantness of an auditory display is the ability to make it silent.

Most perceptual auditory dimensions have a relative low resolution. Still, a lot of information can be communicated by using characteristics of sound efficiently. It is known, and seen during the evaluations, that tempo and pitch have a relatively high resolution. Panning and a crossfade between an original and a highpass filtered sample had a relatively low resolution. The necessary resolution needed for a variable should be assessed and then mapped to a fitting aspect of sound. Care should be taken with using different characteristics

of sound, as they could interact. The lack of orthogonality is a known problem for designing auditory displays. An example of two aspects of sound interacting can be seen in the coordinate-based auditory display. In this auditory display the crossfade with a highpass filtered sample interfered with the change in pitch.

One should also look for intuitive representations of variables. Tempo, for instance, is an intuitive characteristic of sound for representing distance. Panning is an intuitive characteristic for representing left, middle and right. 3D Surround Sound is an intuitive for representing distance and direction. When no intuitive characteristic can be found, one should use a characteristic that possesses easy distinguishable values/levels to which a variable can be mapped. Users should be given time and explanation to learn and internalize this mapping. Because no absolute values are perceived through characteristics of sound it is advisable to present reference values.

All possible events in a virtual environment should be identified and assessed on importance to the user. Providing feedback on important events confirms the actions of the user and makes it easier to understand what is happening. For example, by sonifying the event of coming into contact with a cube, users would less likely identify the wall as a cube.

### **Detecting objects in a virtual environment**

To be able to detect an object in a virtual environment, it should somehow be distinguishable. This can be done in many ways. Sonifying objects separately in time is a good method as demonstrated with the coordinate-based auditory display. The different drum loops are another good example. The drum loops could be distinguished by their rhythm and composition of instruments. In the virtual microphone auditory display, the drum loops propagated in spherical shapes from objects. The distance the drum loops propagate (the size of the spherical shapes around the objects) is important to the functionality of the auditory display. When the distance over which drum loops are carried is large, one can identify objects from afar. On the other hand, more sounds are bound to reach the cursor, which complicates the perception of the individual sounds. Furthermore, the resolution of the distance information will decrease, since the range is spread over a larger distance. A third good example of how objects can be distinguished is the discriminative tempo assigned to a sequence of unique notes in the prioritized notes auditory display. A note played in a fast tempo asynchronous to the tempo of the other notes, which also changed fast in relation to the distance between the cursor and the sonified object, was easily distinguished from the other notes. The blending of notes in the prioritized notes auditory display was intended to give an impression of the amount and position of objects

that were not close to the cursor. Unfortunately, participants were not able to take advantage of this information during the relatively short evaluation. It ought to be seen, whether more experience with the prioritized notes auditory display will change this.

### **Acquiring objects in a virtual environment**

Results from the coordinate-based auditory display show that feedback on the position of an object, relative to the cursor, is necessary to successfully move towards and make contact with an object. The path to an object is determined by the distance and direction from the cursor to this object. Again there are many ways of sonifying the variable representing distance, and the vector representing direction, in an auditory display. See table 1 in the introduction section, for how the different auditory displays sonified the relative distance and relative direction. Both tempo and 3D Surround Sound, as used in the auditory displays, were intuitive measures of distance. Changes in tempo and 3D Surround Sound were easily perceived and associated with relative distance. Since the direction from the cursor to an object is not a single variable, but a vector with three components, the direction is more difficult to sonify. In the nearest path auditory display, the direction vector was represented by three different characteristics of sound. To understand the relative direction the user had to learn eight different sounds (up, center, down; left, center, right; forward combined with center, and backward). Participants were not able to fully internalize these directional cues during the evaluations. The directional cues from the 3D Surround Sound should not take long to understand, since these cues are continuously used during daily life. However, the cues produced by the 3D Surround Sound were not effective enough to represent direction well. More experience of the users, better HRTFs (preferably individualized) and tuning of the auditory display, soundcard and headphone are expected to improve the perception of direction. A completely different way of representing the relative direction to an object, is giving auditory feedback on the orientation of the stylus when it is pointing at an object, as implemented in the pointing-based auditory display. Kinesthetic feedback and auditory feedback are thus combined to represent a direction. Though the principle of pointing is based on visual cues, the visual impaired participants had little problems understanding this auditory display. The pointing principle was by far the most effective way of representing direction during the evaluations.

### **Haptic magnetism and auditory displays**

Haptic magnetism is easy to understand and intuitive to use. It does not give a continuous overview of a virtual environment, but it makes the user find an object when the cursor comes close to it. It is questionable whether haptic magnetism does a good job in representing distance and direction. Users did not take time to feel in which

direction the force was pulling and then move along that direction, but when users felt a force, they let the force drag the cursor to the object. Haptic magnetism was easy to combine with an auditory display. Haptic magnetism is expected to combine well with an auditory display that is good in getting the cursor close to an object, because the magnetism will then pull the cursor towards this object. Care should be taken, however, with implementing haptic magnetism. The forces involved should be rather weak, to limit interference with intended cursor movements and to limit interference with perceiving other forces, for example, object contact forces. As haptic magnetism is only one of several haptic tools for finding objects, also other haptic tools should be considered for combination with an auditory display.

### Some final remarks on the haptic device

During the evaluation it was found that the intended grip of the stylus of the PHANTOM Desktop haptic device is not natural for visual impaired persons. The natural grip was not fully supported by the haptic device, as mechanical forces appeared in a few specific places of the workspace. SensAble technologies, the producer of the PHANTOM Desktop haptic device, also supplies a thimble which is interchangeable for the stylus. A comparison between the thimble and stylus should clarify which one is better suited for use by visual impaired person.

Use of a haptic device can cause strain in the upper arm. One participant reported strain in the upper arm during the user studies. The strain was caused by having to carry the weight of ones arm and some extra weight from the haptic device. A support for the elbow of the arm using the haptic device is expected to solve this potential problem. An extensive research on the ergonomics on the PHANTOM Desktop and comparable haptic devices is advised.

## LITERATURE

Alcañiz, M., Lozano, J., & Rey B. (2004). Technological background about VR. In G. Riva, C. Botella, P. Légeron, & G. Optale (Eds.), *Cybertherapy: Internet and Virtual Reality as Assessment and Rehabilitation Tools for Clinical Psychology and Neuroscience* (pp. 199-214). Amsterdam: Ios Press.

Brewster, S. (1994). Providing a structured method for integrating non-speech audio into human-computer interfaces. Doctoral thesis, University of York, York.

Brewster, S. (2003). Nonspeech auditory output. In J. A. Jacko & A. Sears (Eds.), *The Human Computer Interaction Handbook* (pp. 220-239). Mahwah: Lawrence Erlbaum Associates.

Burdea, G. (2000). Haptic feedback for virtual reality. *Special issue on Virtual Prototyping, International Journal of Design and Innovation Research*, 2(1), 17-29.

Clark, F., & Horch, K. (1986). Kinesthesia. In K. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of Perception and Human Performance v.1 Sensory Processes and Perception* (pp. 13.1-13.62). New York: Wiley.

Edwards, A. (1989). Soundtrack: An Auditory Interface for Blind Users, *Human-Computer Interaction*, 4(1), 45-66.

Edwards, W., Mynatt, E., & Stockton, K. (1994). Providing Access to Graphical User Interfaces — Not Graphical Screens. In *Proceedings of the first annual ACM conference on Assistive technologies, ASSETS '94* (pp. 47-54). New York: ACM Press.

Gaver, W. (1986). Auditory icons: Using sound in computer interfaces. *Human-Computer Interaction*, 2, 167-177.

Gaver, W. (1989). The SonicFinder: An Interface that Uses Auditory Icons, *Human-Computer Interaction*, 4(1), 67-94.

Goldstein, E. (1999). *Sensation and perception* (5th ed.) Chapters 7, 11 & 12. Pacific Grove, CA: Brooks/Cole.

Iwata, H. (2003). Haptic interfaces. In J. A. Jacko & A. Sears (Eds.), *The Human Computer Interaction Handbook* (pp.206-219). Mahwah: Lawrence Erlbaum Associates.

Klatzky, R., & Lederman, S. (1999). The haptic glance: A route to rapid object identification and manipulation. In D. Gopher & A. Koriat (Eds.) *Attention and Performance XVII. Cognitive regulations of performance: Interaction of theory and application* (pp. 165-196). Mahwah: Erlbaum.

Kramer, G. (1994). An introduction to auditory display. In G. Kramer (Ed.), *Auditory Display: Sonification, Audification and Auditory Interfaces* (pp. 1-77). Reading: Addison-Wesley.

Lederman, S., & Klatzky, R. (1987). Hand movements: A window into haptic object recognition. *Cognitive Psychology*, 19, 342-368.

Lemmetty, S. (1999). Review of speech synthesis technology, Masters thesis, Helsinki University of Technology, Espoo.

- MacLean, K. (2000). Designing with haptic feedback. in *Proceedings of IEEE Robotics and Automation ICRA'2000* (pp. 783-788). Retrieved November 15, 2006 from <http://www.cs.ubc.ca/~maclean/publics/icra00-DesignWithHaptic-reprint.PDF>
- Magnusson, C., & Rasmus-Gröhn, K. (2005). Audio haptic navigational tools for non visual environments, In *Proceedings of the first ENACTIVE Workshop*. Retrieved November 15, 2006 from [http://www.certec.lth.se/haptics/papers/ENACTIVE05\\_AudioHapticToole\\_CMagnusson1.pdf](http://www.certec.lth.se/haptics/papers/ENACTIVE05_AudioHapticToole_CMagnusson1.pdf)
- Mereu, S., & Kazman, R. (1996). Audio enhanced 3D interfaces for visually impaired users. In *Human Factors in Computing Systems, CHI'96 Conference Proceedings* (pp 72-78). New York: ACM Press.
- Oakley, I., McGee M., Brewster, S., & Gray, P. (2000) Putting the feel in 'look and feel'. In *Proceedings of the ACM CHI 2000 Human Factors in Computing Systems Conference* (pp. 415-422). New York: ACM Press.
- Pitt, I., & Edwards, A. (1991). Navigating the Interface by Sound for Blind Users. In D. Diaper & N. Hammond (Eds.), *People and Computers VI, Proceedings of the CHI '91 Conference* (pp.373-383). Cambridge University Press.
- Sallnäs, E. (2004). The effect of modality on social presence, presence and performance in collaborative virtual environments, Doctoral thesis, Royal Institute of Technology, Stockholm.
- Savidis, A., Stephanidis, C., Korte, A., Crispian, K., & Fellbaum, K. (1996). A generic direct-manipulation 3D-auditory environment for hierarchical navigation in non-visual interaction. In *Proceedings of the second annual ACM conference on Assistive technologies, ASSETS '96* (pp. 117-123). New York: ACM Press.
- Sekuler, R., & Blake, R. (2002). *Perception* (4th ed.) Chapter 11. New York: McGraw-Hill.
- Sjöström, C. (2001a). Designing Haptic Computer Interfaces for Blind People, In *Proceedings of the sixth international symposium on signal processing and its applications* (pp. 68-71). Malaysia: Universiti Teknologi Malaysia.
- Sjöström, C., (2001b). Virtual Haptic Search Tools - The White Cane in a Haptic Computer Interface, In *Proceedings AAATE 2001, 6th European conference for the advancement of assistive technology* (pp. 68-71). Retrieved November 15, 2006 from <http://www.certec.lth.se/doc/virtualhapticsearch/virtualhapticsearch.pdf>
- Snibbe, S., MacLean, K. Shaw, S., Roderick, J., Verplank, W., & Scheeff, M. (2001). Haptic Techniques for Media Control. In *Proceedings of the 14th annual ACM symposium on user interface software and technology, UIST '01* (pp. 199-208). New York, ACM Press.
- Srinivasan, M., Basdogan, C. (1997). Haptics in Virtual Environments: Taxonomy, Research Status, and Challenges. *Computer and Graphics*, 21(4), 393-404.
- Subramanian, S., Gutwin, C., Sanchez, M.N., Power, C., & Liu, J. (2005). Haptic and Tactile Feedback in Directed Movements. In *Proceedings of GOTH1'05 guidelines on tactile and haptic interactions*. Retrieved November 15, 2006 from [http://www.cs.usask.ca/research/research\\_groups/userlab/GOTH1/Sriram.pdf](http://www.cs.usask.ca/research/research_groups/userlab/GOTH1/Sriram.pdf)
- Tang, A., McLachlan, P., Lowe, K., Saka, C., & MacLean, K. (2005). Perceiving ordinal data haptically under workload, In *Proceedings of the 7th international Conference on Multimodal interfaces, ICMI '05* (pp. 317-324). New York: ACM Press.
- Wall, S., & Brewster, S. (2004). Providing external memory aids in haptic visualisations for blind computer users. In *Proceedings of the Fifth International Conference on Disability, Virtual Reality and Associated Technologies* (pp.157-164), Reading: University of Reading.
- Winberg, F. (2001). Auditory direct manipulation for blind computer users, Licentiate thesis, Royal Institute of Technology, Stockholm.
- Zecca, M., Micera, S., Carozza, M., & Dario, P. (2002). Control of Multifunctional Prosthetic Hands by Processing the Electromyographic Signal, *Critical Reviews™ in Biomedical Engineering*, 20(4-6), 459-485.

## **APPENDIX I**

Procedure of the first user study

## START

## USER STUDY 1.X

- Have the participant sit down in front of the haptic device. (Adjust to Right/Left-handed)
- Start all recording equipment.
  1. WebCam
  2. Audio Recorder
  3. Video Camera
- Produce a Sync Cue for all recordings
- Ask for video permission

“Before we start with this user study, I would like to ask for your permission to record this session. The recordings will only be used for the purpose of this research. Is that okay with you?”

### Introduction

I'll first give you a short introduction on my research. It is a study on how we can use sound to aid visually impaired persons using a haptic device to navigate a virtual 3D environment.

A haptic device is a small robotic arm that produces force feedback when you touch an object in virtual space. One such haptic device is standing in front of you on the table. Attached to the robotic arm is a pen with one button. You can grab this pen like you would grab a normal pen and you can let your index finger rest on the button. By moving the pen around you can explore a virtual environment that is generated by the computer.

Now, I'm going to let you try an example of a haptic virtual environment.

- Start up SimpleEnvironment.exe, 2 Minute exploration

Here you have a room with two large objects in them. Try to find both and discover what shape they have.

---

Okay, you seem to get the hang of it, let's move on. During the test we are going to work with a similar environment. Only the room will be filled with a varying number of small cubes. The cubes can be anywhere in the room. When you touch a cube, you can press the button on the pen, to drag the cube around the room.

From now on we are going to use the headphones. During the study you are going to listen to five different ways of mapping sound to a virtual environment. They might or might not help you in exploring the virtual environment. It is important that you try out each way of presenting sound with an open mind. Try to do the tasks as best as you can. Some may be easy, and others will take some more time. I'm also wearing headphones, so I will hear the same as you do. *(If participant is not blind) Please put on the mask that is on the table, so you can completely concentrate on what you feel and hear.*

- If participant is not blind, put on blindfold.*
- Adjust and put on Headphones

## Questions & Tasks (D: THE PIANO METHOD)

Okay, on to the next method. The **PIANO** method.

This sound mapping I call “the piano method”. Each cube is represented by a different note of a piano. The rhythm in which a note is played indicates how far the tip of your pen is from a cube. The faster the rhythm, the nearer you are. Also when nearing a cube the piano tone will get louder and more clear. You can try out the piano method first with two cubes in the room.

- Start up environment with two cubes (2) → (D)

Listen to how the sound changes when you move around the cube.  
Now try to pick up the cube and move it around.

Now I’m going to start up a new environment and I’ll ask you some questions about it.

- Start up new environment. (R)
  - 1) Can you tell me where most cubes are located?
- Start up new environment. (R)
  - 2) Can you find out how many cubes there are?
- If 2 was easy, start up new environment. (R)*
  - 3) Again, can you find out how many cubes there are?

Great, now I’ll ask you to do some tasks. Try to finish the task as fast as you can. I’ll start up a new environment. This time there will always be five cubes.

- Start up new environment. (5)
  - 1) Move all the cubes to the **bottom** wall of the room, you may start... NOW.
- If 1 was easy, start up new environment. (5)*
  - 2) Move 3 cubes to the **front** wall and 2 cubes to the **back** wall, you may start... NOW.
- Start up new environment. (5)
  - 3) Place all the cubes on top of each other in the center of the floor, you may start... NOW.

## Reflection

1. What did you think of the **PIANO** audio method?
2. Did you find the questions that I asked in the beginning easy or difficult?
3. Can you tell me why the questions were easy or difficult?
4. Did you find the tasks I gave you easy or difficult?
5. Can you tell me why the tasks were easy or difficult?
6. Did you like how this method sounded? Was it nice to listen to or annoying?
7. Could you use it for a long time?
8. How would you improve this **PIANO** method?



## Questions & Tasks (A: THE SCANNING METHOD)

Okay, on to the next method. The SCANNING method.

You can hear a continuous low pitched tone moving from the left to right over and over. The horizontal position of the sound indicates the position of the scanner. When the scanner meets a cube, this cube will make a bell sound. The properties of this bell sound depend on the position of the cube. A high pitched bell sound means that the cube is located near the top of the room, and a low pitched bell sound means that the cube is located near the bottom of the room. When a bell sound comes from the right headphone-speaker, the cube is located to the right of the room, and when a bell sound comes from the left headphone-speaker the cube is located to the left of the room. When a bell sound is muffled the cube is located to the back of the room and when the bell sound is clear the cube is located to the front of the room. You can try out the scanning method first with two cubes in the room.

- Start up environment with two cubes (2) → (A)

Listen to how the sound changes when you move around the cube.  
Now try to pick up the cube and move it around.

Now I'm going to start up a new environment and I'll ask you some questions about it.

- Start up new environment. (R)
  - 4) Can you tell me where most cubes are located?
- Start up new environment. (R)
  - 5) Can you find out how many cubes there are?
- If 2 was easy, start up new environment. (R)*
  - 6) Again, can you find out how many cubes there are?

Great, now I'll ask you to do some tasks. Try to finish the task as fast as you can. I'll start up a new environment. This time there will always be five cubes.

- Start up new environment. (5)
  - 4) Move all the cubes to the **back** wall of the room, you may start... NOW.
- If 1 was easy, start up new environment. (5)*
  - 5) Move 3 cubes to the **upper** wall and 2 cubes to the **bottom** wall, you may start... NOW.
- Start up new environment. (5)
  - 6) Place all the cubes on top of each other in the center of the floor, you may start... NOW.

### Reflection

9. What did you think of the SCANNING method?
10. Did you find the questions that I asked in the beginning easy or difficult?
11. Can you tell me why the questions were easy or difficult?
12. Did you find the tasks I gave you easy or difficult?
13. Can you tell me why the tasks were easy or difficult?
14. Did you like how this method sounded? Was it nice to listen to or annoying?
15. Could you use it for a long time?
16. How would you improve this SCANNING method?

## Questions & Tasks (B: THE SWITCHING METHOD)

We will start with “the SWITCHING method”.

This sound mapping I call “the single method”. When you move the pen around you will hear a bell sound being played over and over, but with changing rhythm and other properties. The bell sound that you hear represents the closest cube to the tip of your pen. The faster the rhythm of the bell sound the closer you are to a cube. The properties of the bell sound give you directions to the cube. When the bell sound has a high pitch the cube is located higher than the tip of your pen, and when the bell sound has a low pitch the cube is located lower than the tip of your pen. When the bell sound comes from the left, the cube is to the left of the tip of your pen, and when from the right the cube is right of the tip of your pen. When the bell sound is muffled it means that the tip of the pen is behind the cube. When you move around you will sometimes also hear a short ‘boupboup’ sound, this means you have moved closer to another cube. And you are hearing directions now to find that other cube. You can try out the single method first with two cubes in the room.

- Start up environment with two cubes (2) → (B)

Listen to how the sound changes when you move around the cube.  
Now try to pick up the cube and move it around.

Now I’m going to start up a new environment and I’ll ask you some questions about it.

- Start up new environment. (R)
  - 1) Can you tell me where most cubes are located?
- Start up new environment. (R)
  - 2) Can you find out how many cubes there are?
- If 2 was easy, start up new environment. (R)*
  - 3) Again, can you find out how many cubes there are?

Great, now I’ll ask you to do some tasks. Try to finish the task as fast as you can. I’ll start up a new environment. This time there will always be five cubes.

- Start up new environment. (5)
  - 1) Move all the cubes to the **front** wall of the room, you may start... NOW.
- If 1 was easy, start up new environment. (5)*
  - 2) Move 3 cubes to the **left** wall and 2 cubes to the **right** wall, you may start... NOW.
- Start up new environment. (5)
  - 3) Place all the cubes on top of each other in the center of the floor, you may start... NOW.

### Reflection

1. What did you think of the SWITCHING method?
2. Did you find the questions that I asked in the beginning easy or difficult?
3. Can you tell me why the questions were easy or difficult?
4. Did you find the tasks I gave you easy or difficult?
5. Can you tell me why the tasks were easy or difficult?
6. Did you like how this method sounded? Was it nice to listen to or annoying?
7. Could you use it for a long time?
8. How would you improve this SWITCHING method?

#####  
##### 10 MINUTES BREAK #####  
#####

## Questions & Tasks (E: THE POINTING METHOD)

Okay, on to the next method. The **POINTING** method.

This method uses the pen as a pointing device. When you point the pen towards a cube, it will make a bell sound over and over. The more near you get to a cube the faster the rhythm of these bells. When you point, not directly at, but close to a cube, a guiding tone can be heard. A high guiding tone indicates that you should rotate the pen upwards, and with a low tone, downwards. When the guiding tone is coming from the left you should rotate the pen to the left and when the guiding tone is coming from the right you should rotate the pen to the right. You can try out the pointing method first with two cubes in the room.

- Start up environment with two cubes (2) → (E)

Listen to how the sound changes when you move around the cube.  
Now try to pick up the cube and move it around.

Now I'm going to start up a new environment and I'll ask you some questions about it.

- Start up new environment. (R)
  - 7) Can you tell me where most cubes are located?
- Start up new environment. (R)
  - 8) Can you find out how many cubes there are?
- If 2 was easy, start up new environment. (R)*
  - 9) Again, can you find out how many cubes there are?

Great, now I'll ask you to do some tasks. Try to finish the task as fast as you can. I'll start up a new environment. This time there will always be five cubes.

- Start up new environment. (5)
  - 7) Move all the cubes to the **left** wall of the room, you may start... NOW.
- If 1 was easy, start up new environment. (5)*
  - 8) Move 3 cubes to the **right** wall and 2 cubes to the **left** wall, you may start... NOW.
- Start up new environment. (5)
  - 9) Place all the cubes on top of each other in the center of the floor, you may start... NOW.

## Reflection

17. What did you think of the **POINTING** method?
18. Did you find the questions that I asked in the beginning easy or difficult?
19. Can you tell me why the questions were easy or difficult?
20. Did you find the tasks I gave you easy or difficult?
21. Can you tell me why the tasks were easy or difficult?
22. Did you like how this method sounded? Was it nice to listen to or annoying?
23. Could you use it for a long time?
24. How would you improve this **POINTING** method?

## Questions & Tasks (C: THE MUSIC METHOD)

Okay, on to the next method. The **MUSIC** method.

Each cube is represented by a different music loop. The tip of your pen now acts like a microphone through which you listen to the room. When you are near a cube it will sound loud and clear and when further away it will sound soft and distant. If you listen carefully you should also be able to hear when you are in front, back, above or below of the cube. You can try out the music method first with two cubes in the room.

- Start up environment with two cubes (2) → (C)

Listen to how the sound changes when you move around the cube.  
Now try to pick up the cube and move it around.

Now I'm going to start up a new environment and I'll ask you some questions about it.

- Start up new environment. (R)
  - 10) Can you tell me where most cubes are located?
- Start up new environment. (R)
  - 11) Can you find out how many cubes there are?
- If 2 was easy, start up new environment. (R)*
  - 12) Again, can you find out how many cubes there are?

Great, now I'll ask you to do some tasks. Try to finish the task as fast as you can. I'll start up a new environment. This time there will always be five cubes.

- Start up new environment. (5)
  - 10) Move all the cubes to the **front** wall of the room, you may start... NOW.
- If 1 was easy, start up new environment. (5)*
  - 11) Move 3 cubes to the **upper** wall and 2 cubes to the **bottom** wall, you may start... NOW.
- Start up new environment. (5)
  - 12) Place all the cubes on top of each other in the center of the floor, you may start... NOW.

## Reflection

25. What did you think of the **MUSIC** method?
26. Did you find the questions that I asked in the beginning easy or difficult?
27. Can you tell me why the questions were easy or difficult?
28. Did you find the tasks I gave you easy or difficult?
29. Can you tell me why the tasks were easy or difficult?
30. Did you like how this method sounded? Was it nice to listen to or annoying?
31. Could you use it for a long time?
32. How would you improve this **MUSIC** method?

## Audio Mode Comparison

Now we are going to compare the different sound mappings. I want you to rank the five different mappings first on usefulness and thereafter on pleasantness.

So first rank the sound mappings on usefulness. The sound mapping that you find the most useful gets rank 1. The sound mapping that you find the least useful gets rank 5. You can listen to all sound mappings again in any order. You can tell me which one you want to hear. I'll first go through all five sound mappings as you heard them in this test.

- Start up new environment with 5 cubes (5)
- Go through all the audio modes. (D) → (A) → (B) → (E) → (C)

Usefulness		
1.	PIANO	
2.	SCANNING	
3.	SWITCHING	
4.	POINTING	
5.	MUSIC	

Now I want you to rank the five different mappings on pleasantness. The sound mapping that you liked listening to most gets rank 1. The sound mapping that you liked listening to the least gets rank 5. You can listen to all sound mappings again in any order. You can tell me which one you want to hear. I'll first go through all five sound mappings as you heard them in this test.

- Go through all the audio modes. (D) → (A) → (B) → (E) → (C)

Pleasantness		
1.	PIANO	
2.	SCANNING	
3.	SWITCHING	
4.	POINTING	
5.	MUSIC	

## Magnetism Exploration

Now we are going to try out something new. I'm going to start a new environment again, but now all the cubes will be magnetic. The tip of the pen will be attracted to the cube if it is close enough. You can explore how that feels first.

- Start up new environment with two cubes, enable gravity wells (2) → (G)

Okay now I'm going to ask you questions again, just like I did before with the different sound mappings.

- Start up new environment. (R)
  - 1) Can you tell me where most cubes are located?
- Start up new environment. (R)
  - 2) Can you find out me how many cubes there are?
- Start up new environment. (5)
  - 1) Place all five cubes on top of each other in the center of the floor, you may start... NOW.

R1) How did you like the magnetism? Did you find the questions tasks easy/hard?

Now I'm going to enable the **POINTING** method, and we'll go trough the same questions.

- Start up new environment. (R) → (E)
  - 1) Can you tell me where most cubes are located?
- Start up new environment. (R)
  - 2) Can you find out me how many cubes there are?
- Start up new environment. (5)
  - 1) Place all five cubes on top of each other in the center of the floor, you may start... NOW.

R2) How did you like it with the **pointing** method? Did you find the questions tasks easy/hard?

Now I'm going to enable the **MUSIC** method, and we'll go trough the same questions.

- Start up new environment. (R) → (C)
  - 1) Can you tell me where most cubes are located?
- Start up new environment. (R)
  - 2) Can you find out me how many cubes there are?
- Start up new environment. (5)
  - 1) Place all five cubes on top of each other in the center of the floor, you may start... NOW.

R3) How did you like it with the **music** method? Did you find the questions tasks easy/hard?

Now I'm going to enable the **SCANNING** method, and we'll go trough the same questions.

- Start up new environment. (R) → (A)
  - 1) Can you tell me where most cubes are located?
- Start up new environment. (R)
  - 2) Can you find out me how many cubes there are?
- Start up new environment. (5)
  - 1) Place all five cubes on top of each other in the center of the floor, you may start... NOW.

R4) How did you like it with the **scanning** method? Did you find the questions tasks easy/hard?

**THE END.**

## **APPENDIX II**

### Short Paper

Presented by Fredrik Winberg at

*The Eighth International ACM SIGACCESS Conference on Computers & Accessibility.*  
Portland, USA, October 2006.

# Designing Auditory Displays to Facilitate Object Localization in Virtual Haptic 3D Environments

**Koen Crommentuijn**

Department of Human-Computer Interaction  
School of Computer Science and Communication  
Royal Institute of Technology  
SE-100 44 Stockholm, Sweden  
kjcr@csc.kth.se

**Fredrik Winberg**

Department of Human-Computer Interaction  
School of Computer Science and Communication  
Royal Institute of Technology  
SE-100 44 Stockholm, Sweden  
fredrikw@csc.kth.se

## ABSTRACT

Five different auditory displays were designed to aid blind users in finding objects in a virtual haptic 3d environment. Each auditory display was based on a different principle and incorporated different methods for representing spatial information. Results from an evaluation with seven visually impaired persons reveal to what extent these methods facilitate object localization in a virtual haptic 3d environment.

## Categories and Subject Descriptors

H.5.2 [Information Systems]: User Interfaces - *Auditory (non-speech) feedback, Haptic I/O*

## General Terms

Design, Human Factors.

## Keywords

Auditory display, haptics, object localization, visually impaired users.

## 1. INTRODUCTION

With the development of haptic devices, such as the PHANTOM Desktop<sup>1</sup>, visually impaired persons have gained the means to explore virtual 3d environments. But when no visual representation is used, the task of finding objects in virtual space can become time consuming and frustrating. Every inch has to be scanned in order to know what is there and when a virtual environment is dynamic, many events and changes will go unnoticed to the user.

Researchers and software developers within the haptic community have proposed a variety of haptic tools to aid users in localizing objects (e.g. [1, 5]). The use of an auditory display for the same purpose has received less attention. Because so far only a few auditory displays have been investigated, this research set out to implement and compare new methods to facilitate object localization in virtual haptic 3d environments.

## 2. AUDITORY DISPLAYS

An auditory display uses non-speech sound to convey information to a user. In this case the information should represent or lead to the position of objects in a virtual 3d environment. One way is to

Another way is, to sonify a virtual environment in relation to the position of the cursor (point of interaction). From the cursor the direction and/or distance to objects can be sonified. The second of the five auditory displays investigated uses discrete cues (e.g. up-right-front) to represent direction combined with an increasing rhythm when closing in on an object to represent distance. The third auditory display uses 3d surround sound to create a natural audio-haptic virtual environment. Like Pitt & Edwards [4], the cursor here acts as a virtual microphone through which one can listen to the continuously playing virtual sound sources placed in virtual space. The fourth auditory display uses 3d surround sound to (primarily) represent direction combined with an increasing rhythm when closing in on an object to represent distance.

Yet another way to sonify a virtual environment is to use the direction in which the cursor is pointing (cf. [3]). The direction to an object is thus represented by the position of the pen (kinesthetic) in combination with audio feedback. The fifth auditory display is based on this principle and also incorporates a guiding tone when pointing near to an object to help the user to point directly at the object. The distance is again represented by an increasing rhythm.

Haptic magnetism, a haptic aid to find objects, was chosen to be implemented along with the auditory displays to evaluate how a haptic aid works. With haptic magnetism enabled, objects exert small attractive forces upon the haptic device, pulling the hand of a user towards the objects.

## 3. EVALUATION

All audio methods have been implemented in a 'simple' virtual haptic 3d environment (see figure 1). The environment was bounded by a large box in which three to eight cubes were present.

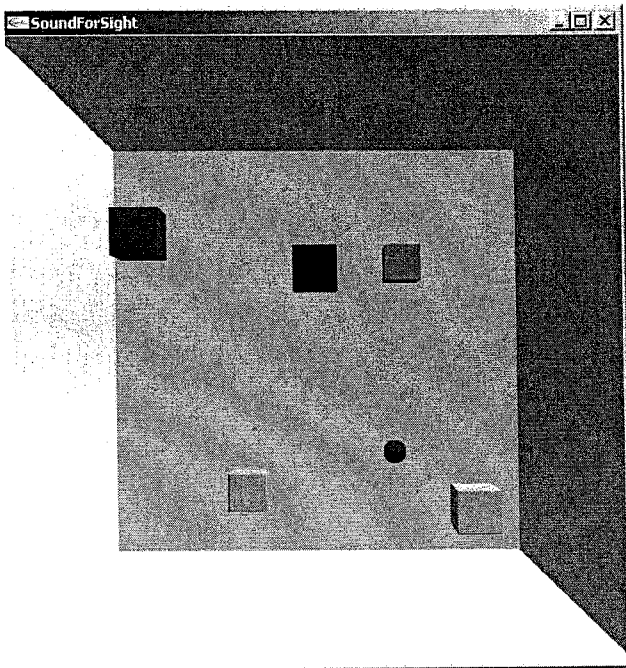
The virtual environment could be explored with a PHANTOM Desktop. When touching a cube and simultaneously pressing the button on the stylus, the cube could be dragged. Audio was played over a pair of headphones. No visual representation was given to the user.

The evaluations involved questions about the number of objects and the layout of the virtual environment as well as tasks to move the cubes into specific positions (e.g. move all cubes to the left

---

<sup>1</sup> [www.sensable.com](http://www.sensable.com)





**Figure 1: Visual representation of the virtual environment.**

wall). The moving tasks provided the users with an interesting challenge which could easily be timed. Three visually impaired persons (2 male, 1 female) tried all 5 auditory displays during an evaluation. Four visually impaired persons (2 male, 2 female) tried only 2 pre-selected auditory displays, to obtain a more detailed evaluation. Except one female who had low vision (10/100), all participants were blind. The age of the participants varied from 26 to 59.

## 4. RESULTS

### 4.1 Coordinate Sonification

Sonifying the x,y,z-coordinates of objects did not successfully aid users in finding them. The characteristics of sound (pitch, panning, high-pass filter) that were mapped to the position of an object lack resolution and don't provide absolute values. The implemented auditory display did give the users a global impression, but left a too large area to search for the objects.

### 4.2 Discrete Directional Cues

Using discrete cues to represent the direction of objects is useful. It takes some time for a new user to recognize and make use of the cues, but they then provide an unambiguous direction. Integrating discrete directional cues into an auditory display can be difficult though. An environment with more than a few objects needs a well designed algorithm in order to present all the discrete directions without confusing the user.

### 4.3 Distance by Rhythm

An increasing rhythm to represent the distance between the cursor and an object appeared intuitive. Users easily perceived when they moved the cursor nearer to an object. Using rhythm

to represent distance can be a valuable option when combining it with another method to represent direction.

### 4.4 Virtual Microphone

Positioning virtual sound sources at the location of virtual objects creates a natural/realistic auditory space. The use of such a method is very intuitive and requires little cognitive effort by the user. The time to find-and-touch objects was on average the shortest with the implementation of this auditory display. It should be noted however that the fine movement necessary to go from being close to touching an object took a relative long time. Providing more support for the final fine movement should improve the performance of this auditory display.

### 4.5 Pointing

For a visually impaired person, the principle of pointing is less obvious, because it is based on a visual cue. But the results show that with some practice and explanation the users were able to successfully use the stylus of the PHANTOM Desktop as a pointing device. It took some time and 'luck' to discover all the objects present, but when an object was discovered, users had little trouble to move closer and touch it.

### 4.6 Haptic Magnetism

The users reacted positively to haptic magnetism. Some tasks were completed amazingly fast, but some similar tasks took a long time to complete or were not completed correctly. It was reported that haptic magnetism, despite the attractive forces being weak, from time to time interfered with the desired movement of the user. Users also reported that when they felt the haptic magnetism, they let it pull the stylus toward the virtual object. The haptic magnetism is thus not used to extract the direction or distance to an object, but used in a more passive way. Since haptic magnetism is only felt within a certain range of an object, a supporting method giving a more global overview would be useful.

## 5. REFERENCES

- [1] Magnusson, C. & Rasmus-Gröhn, K. *Audio haptic navigational tools for non visual environments*, First ENACTIVE Workshop, Pisa, Italy, 2005.
- [2] Mereu, S. & Kazman, R. *Audio Enhanced 3D Interfaces for Visually Impaired Users*, Proceedings of CHI '96, Vancouver, Canada, 1996.
- [3] Müller-Tomfelde, C. *Interaction Sound Feedback in a Haptic Virtual Environment to Improve Motor Skill Acquisition*, Proceedings of ICAD '04, Sydney, Australia, 2004.
- [4] Pitt, I. & Edwards, A. *Navigating the Interface by Sound for Blind Users*, In D. Diaper & N. Hammond (Eds.), *People and Computers VI*, Proceedings of the CHI '91 Conference (pp. 373-383). Cambridge University Press, 1991.
- [5] Sjöström, C. *Virtual Haptic Search Tools – The White Cane in a Haptic Computer Interface*, Presented at AAATE 2001, Ljubljana, Slovenia, 2001.