

# Audio makes a difference in haptic collaborative virtual environments

JONAS MOLL, YING YING HUANG, EVA-LOTTA SALLNÄS

HCI Dept., School of Computer Science and Communication, Royal Institute of Technology, Sweden

## **Abstract**

In this paper a study is presented which aimed at exploring the effects of audio feedback in a haptic and visual interface supporting collaboration among sighted and people who cannot see. A between group design was used and the participants worked in pairs with one sighted and one blindfolded in each. The application used was a haptic 3D environment in which participants could build composed objects out of building blocks. The building blocks could be picked up and moved around by means of a touch feedback pointing device. In one version of the application sound cues could be used to tell the other person where you were, and to get feedback on your own and the other person's actions. Results showed that sound cues together with haptic feedback made a difference in the interaction between the collaborators regarding their shared understanding of the workspace and the work process. Especially, sound cues played an important role for maintaining awareness of ongoing work – you knew what was going on, and you got a response on your own actions.

**Keywords:** Haptic, Audio, Multimodal Interfaces, Collaboration, Problem solving

# 1. Introduction

Collaboration becomes increasingly important in educational environments. Especially in school, the pupils collaborate in solving all sorts of tasks together as a way of learning through social interaction. If one of the collaborating pupils is visually impaired this might cause difficulties since the most important sense – vision – is not available. During the last few years attempts have been made to develop different software that supports visually impaired pupils in school. However, due to the fact that these tools are mostly designed for individual work they do not work very well in a group work situation.

In an earlier study (<reference to own work>) a collaborative application, aiming at supporting pupils in learning about spatial geometry, was developed and evaluated. In that application pupils could feel the shape of different geometrical objects, which could be picked up and moved around using a touch feedback system. The pupils could build composed objects out of smaller ones. The application was evaluated in pairs consisting of sighted and visually impaired children in four elementary schools in the Stockholm area. The results shed light on several important aspects of haptic feedback when it came to support for collaboration between sighted and visually impaired children. The need for reference points was highlighted and we could see that the haptic feedback had a positive effect on the inclusion of the visually impaired pupil and that it could be an aid when discussing strategies and geometrical objects. Haptic guiding functions, by which the sighted pupil could drag the visually impaired pupil around was also shown to be a great aid. However, the results from the evaluation also showed that it was problematic to maintain awareness in this kind of dynamic environment where all users could move objects around.

Based on the findings from the previous evaluation, four different audio cues were added to the interface and the resulting application was evaluated in the study presented in this paper. The aim of our current study has been to alleviate the problems the visually impaired pupils had in maintaining awareness of other's actions in a shared virtual work space and to make the grounding process easier as well as increasing the sense of social presence.

## 2. Multimodal collaborative interfaces

### 2.1 The issue of equivalent interfaces

The main benefit of working co-located compared to working distributed is that all group members have a good awareness of the presence of others, their activities, use of resources, knowledge, expectations and current goals (Neale et al., 2004). Dourish and Bellotti (1992) define awareness as “an understanding of the activities of others, which provides a context for your own activity”. When collaborating in groups, the level of awareness of other’s activities is crucial (Gaver et al., 1991). Successful collaboration depends on the ability to attend to and understand information produced by other members of the collaborating team – the more feedback one get about other’s actions the more time one can focus on making the collaboration work. Verbal communication and perception of common objects, like physical artefacts, are two aspects that are essential in order to coordinate joint activities for groups and individuals that cooperate (Malone and Crowston, 1990). According to Kraut et al (1993) lack of this type of information decreases the quality in joint projects.

When sight is not available, awareness information cannot be obtained through the visual modality. This often limits a visually impaired person’s ability to obtain awareness. Several attempts have been made to get around this problem by e.g. using Braille displays and screen readers (Mynatt and Weber, 1994). The problem is, however, that the sighted and visually impaired pupils often have access to different work materials and different information and that it is not always easy to translate between the different representations. We argue that even if an exact translation of a GUI is not possible using other modalities like touch or hearing, it is possible to design equivalent interfaces. That means interfaces that include all important functions and basic elements and that give information about actions in the interface. Clark and Brennan (1991) define common ground as a state of mutual understanding among conversational participants about the topic at hand. Furthermore, Clark’s theory of common ground says that people must have shared awareness in order to carry out any form of joint activities (Clark, 1996).

Grounding activities aim to provide mechanisms that enable people to establish and maintain common ground (McCarthy et al., 1991). The use of gestures has been identified as important in the grounding process (Kirk et al., 2007), and here visually impaired people have a clear disadvantage. It is of utmost importance to support the grounding process between a sighted and visually impaired user when developing collaborative applications with these users in mind.

One interesting type of gestures that has been shown to be an important aid in the grounding process is deictic references to objects like e.g., “this one”, “the blue block” and “that wall” . With references like these you direct your partner’s attention to a specific object which both are aware of verbally, by eye gaze or by pointing. Being able to use deictic referencing makes it easier to maintain common ground (Burke and Murphy, 2007). We argue that haptic feedback has communicative properties that can be used by themselves or together with auditory feedback in order to make references to directions and objects.

## 2.2 Haptic and auditory feedback

Not many studies have been conducted in which collaboration is investigated between visually impaired and sighted users in interfaces that provide both auditory and haptic feedback. Haptic perception is a combination of tactile perception (through the skin) and kinaesthetic perception (the position and movement of joints and limbs) (Loomis and Lederman, 1986). An important aspect of haptic perception is that it is mostly obtained by actively exploring objects with the fingers and hands.

Users with a severe visual impairment have to work without the information obtained from vision, which makes task performance more difficult. It is especially hard to get an overview of an interface and to find and explore objects, as well as finding interesting detailed parts of specific objects (Jansson and Juhasz, 2007). Designers are beginning to realize that haptic displays can help blind individuals to overcome the challenges experienced when accessing and exploring the web (Kuber et al., 2007). The touch modality has been shown to make it possible for visually impaired to explore and navigate in virtual environments (Sjöström, 2001). By using the sense of touch visually impaired users can identify and perceive the shape and texture of objects and this enriches the interaction. In a recent EU project with the aim of developing a haptic display for exploration of virtual copies of statues, the art at museums was made accessible to visually impaired people (Bergamasco et al., 2001; Bergamasco and Prisco, 1998; Frisoli et al., 2002).

There are many kinds of devices that give haptic feedback to the user. In this study PHANTOMS, a kind of one-point-interaction devices with three degrees of freedom have been used. A PHANTOM has a pen-like handle, which is attached to a robotic arm. When the virtual tip of the pen “touches” a virtual object forces are generated through the robotic arm, giving the user the impression of actually touching the object.

It has been shown that haptic feedback increases the performance as well as the perceived presence when groups of sighted users solve tasks together (Oakley et al., 2001; Basdogan et al., 2000; <Reference to own work>). In one study, interaction between sighted and visually impaired adult persons was investigated when using three different collaborative haptic interfaces (<reference to own work>). Findings from that study showed that visually impaired and sighted persons could get a shared understanding of the layout of the interface and that the pair could hand off objects to each other and discriminate between objects with different shapes, sizes and softness in a haptic interface. Different guiding strategies used by visually impaired and sighted adults that collaborated, was also investigated (<reference to own work>). A haptic guiding function that allowed one person to guide the other user by grabbing his/her avatar (by using a haptic feedback device) was shown to be useful, especially as a complement to verbal guiding. Verbal guiding was however shown to be the most important kind of guiding, i.e. talking about directions like “go left,... stop, now go down,...”. Since the ability to move objects could create an interesting active learning environment and the benefit of using haptic guidance was a clear aid for the visually impaired adults these functions were further investigated, although in the new context of a problem solving environment, in the study presented in this paper.

Within the European project MICOLE (Multimodal Collaboration Environment for

Inclusion of Visually Impaired Children) several interfaces based on vision, touch and audio for collaboration between sighted and visually impaired children have been developed, although not all of them have been evaluated in a collaborative context. Examples of interfaces are active exploration of simple astronomical phenomena (Saarinen et al., 2005), exploration of electric circuits (Pietrzak et al., 2007) and an audio-haptic drawing application (Rasmus-Gröhn et al., 2006). In all these examples both haptic and audio feedback is given to provide a whole multimodal experience.

Although a number of projects have investigated how interaction in auditory interfaces can be provided to visually impaired users (Poll and Eggen, 1996; Kennel, 1996; Zhao et al., 2008), less attention has been paid to the impact of auditory feedback in combination with haptic feedback in a collaborative setting (apart from the work within the MICOLE project referenced above). Winberg and Hellström (2001) have developed an interface based solely on audio feedback. They invented a sound model that made it possible for blind users to play the popular game Towers of Hanoi. They used either three or four disks and each disk had a unique sound differing in pitch and timbre. The height of a particular disk was represented by the length of the sound and stereo panning was used to convey information about which peg a particular disk was on. The Towers of Hanoi application was tested with groups of sighted and blind adults (Winberg and Bowers, 2004). The sighted person used a visual interface and the blind the auditory interface described above. They had to take turn in moving the disks and they did not have access to each other's representations. Since all pairs managed to solve the game, this shows that it is possible to collaborate even if one of the users only had access to a sound interface.

An auditory and haptic interface in which a set of objects was represented in five different ways was evaluated with visually impaired adults in one study (Crommentuijn, 2006). It was shown that the design in which the user could "hold a virtual microphone" and move around until objects were found was the most efficient. From the results in these studies it can be concluded that information such as the location of objects and even the location and action of your partner in a collaborative haptic context could be represented and conveyed by auditory cues.

Wall and Brewster (2006) have also investigated how audio and haptic feedback in combination can be beneficial for visually impaired users by developing and evaluating haptic and audio bar charts. Results have indicated that visually impaired users can scan bar charts by using haptic feedback provided that there are enough reference points in the interface. Sound was also shown to be an aid searching for more detailed information, something that was also shown by Yu and Brewster (2003). The benefit of using audio feedback through earcons in combination with haptic feedback has also been highlighted in studies of crossmodal icons for mobile devices (Hoggan and Brewster, 2007), where it was also shown that it is possible to identify information given in one modality after being trained in another. Earcons are a kind of non-verbal audio icons that give the status of objects, operations or actions in an interface.

Less attention has been paid to the impact of auditory feedback in combination with haptic feedback in a collaborative setting between sighted and visually impaired users. In one study, the interaction between visually impaired pupils and their teachers was investigated regarding the effects of training handwriting using haptic and audio output to realize a teacher's pen input to the pupil (Plimmer et al., 2008). The researchers observed improvements in the character shapes drawn by especially the completely blind children.

The aim of the study presented in this paper is to fill a gap by focusing on collaboration between sighted and visually impaired people, as well as on interaction in a shared virtual environment that provides both visual, haptic and auditory information.

### 3. The study

A study of collaboration between sighted adults of which one in each pair was blindfolded has been conducted in order to investigate the joint actions in a haptic, auditory and visual interface.

#### 3.1 Participants

A total of 32 participants, in the age span of 25-35 years, took part in the study. The participants were divided into pairs, with one blindfolded and one sighted person in each. The person that was going to be blindfolded was randomly selected. To make sure collaboration was encouraged everyone who wanted to take part in the study got to choose a person they knew, to work with.

Visually impaired people were not recruited for this study, even though it would have been better than blindfolding sighted people. More participants were needed, than could be found in the target group consisting of visually impaired people. Visually impaired people's mental model of space may differ from that of blindfolded sighted people's and that has to be taken into account when generalising the results from this study. However, since a very similar application has been tested with groups of sighted and severely visually impaired (half of these congenitally blind) school children already (see <reference to own work>) it is possible to make a fairly good judgement of validity. The difference between this and the earlier study is that some usability issues have been fixed and that sound cues have been added. The results from the earlier study showed that visually impaired pupils could efficiently build a mental representation of the environment based on haptic feedback but that they had some difficulty in tracking changes e.g. if objects were moved. The hypothesis is that the audio cues in this study can solve that problem. Even though there might be a difference between visually impaired persons and blindfolded sighted persons regarding interpretation of 3D sounds we will still be able to draw important qualitative conclusions regarding the positive effects that the sound has.

#### 3.2 Software and hardware

##### 3.2.1 *The application*

The application investigated in this study is a three-dimensional haptic, auditory and visual virtual environment (figure 2-3). The scene is a room with walls and a floor that have different and discriminable textures applied to them that can be felt using a haptic device. The environment

contains a number of cubic and rectangular building blocks and their shape and surface friction can also be felt. In the environment the users are represented by a red and a blue sphere, respectively.

Apart from feeling and recognizing different geometrical shapes a user can also pick up and move around the objects by means of the PHANTOM. An object is grasped when a user pushes the phantom button when being in contact with the object and dropped when the user releases the button. If two users grasp the same object they feel each other's forces on it. For example, this enables a sighted user to guide a visually impaired (or blindfolded) one to a certain place. Since gravity is applied to all the objects the users feel the weight and inertia as they carry objects around.

Users can also feel and grasp each other's graphical representations in order to provide haptic navigational guidance. They can also "feel each others' proxies" by means of a small repelling force, applied whenever the users' graphical representations get in close distance to each other.

Two different versions of the application were used in this study. In the first version the haptic and visual interface described above was used. In the second one, four types of audio cues were implemented as a complement to the visual and haptic feedback given in the first version. The first type of sound cue, a *grip sound*, is heard every time one lifts an object. The second type is a kind of *touch down sound*, which is heard every time an object touches the floor. The third type is a *collision sound*, heard every time an object lands on top of another. The last type is a *contact sound*, which is heard every time a user pushes the button on the PHANTOM. This 3D contact sound is only heard relative to the blindfolded's avatar (no matter who pushes the button) making it possible for the blindfolded participant to locate the other user's position relative to his/her own position.

### 3.2.2 Apparatus

In the study, we used a Dell Precision PC with two dual core processors. Two PHANTOMs, one PHANTOM Desktop and one PHANTOM Omni, were connected serially to the computer. One ordinary computer screen, a keyboard to load assignments and a mouse to start the application were also connected. The Reachin API 4.1 software, as well as Visual Studio .NET 2003, are needed for the application to run. Stereo glasses were not used in this study and thus the sighted users were not provided depth perception. Camstudio was used for screen capturing during the experimental sessions.

## 3.3 Procedure

A between group design was used in the study comparing a visual/haptic/auditory and a visual/haptic interface. The test was divided into four parts; one demo session for the blindfolded participant, one training session for both participants, one group work session and one interview session. First of all, the researchers gave introductory information about the aim of

the experiment, followed by instructions on how to use the haptic devices. The participant who was going to be blindfolded then got the opportunity to work with a demo program for some minutes, in which the user could feel different textures and surfaces applied to several cubes.

Next, the participants worked together in a training session in which they practiced how to feel the shape of a cube, how to navigate in the three-dimensional environment and how to grab a cube, lift it and hand it off to the other person in the group. The groups that were to use the visual/haptic/auditory interface were also presented with the different kinds of sound. We made sure, before the real tasks were loaded, that the participants felt comfortable with one being blindfolded and that they understood how they could interact with each other and with the objects in the virtual environment.

After the training session, which lasted for about 15 minutes, the participants solved two tasks, described in more detail in the next section. The groups were randomly assigned to either the visual/haptic/auditory or the visual/haptic interface. The blindfolded participant used the PHANTOM Desktop that has better resolution and the sighted one used the PHANTOM Omni. The setup used in the training session and during the experimental tasks is shown in figure 1.



*Fig. 1. Experimental setup. The two participants (the one closest to the camera is blindfolded) are seen holding one PHANTOM device each.*

When the tasks were completed the pairs of participants were interviewed together. The interview was semi structured and lasted for about 20 minutes. The interview questions aimed at exploring the participants' thoughts about usability aspects regarding haptic and audio feedback. Among other things the focus was on what kind of sound cues the participants in the haptic only group would like and if the participants in the audio groups could make use of the audio cues implemented. Questions about common ground, awareness and social presence were also asked to give more insight in how audio might affect these aspects.

### 3.4 Tasks

During the test session the participants solved two tasks in the virtual environments collaboratively.

#### Task 1.

*In this assignment eight cubes of size  $2 \times 2 \text{ cm}^2$  are placed on the floor, four at the bottom left and four at the upper right side of the room respectively. In the middle of the room there is also a large board with a size of  $12 \times 6 \text{ cm}^2$ . Your assignment is to build a table. The task has been solved when the board has been positioned 4 cm above floor level. The table legs should be at the respective corners of the table so as to give a good-looking impression. Before you start to solve this task you have to decide who is responsible for which group of cubes, you only have the right to touch your own cubes.*

Figure 2 shows a screen shot of this assignment.

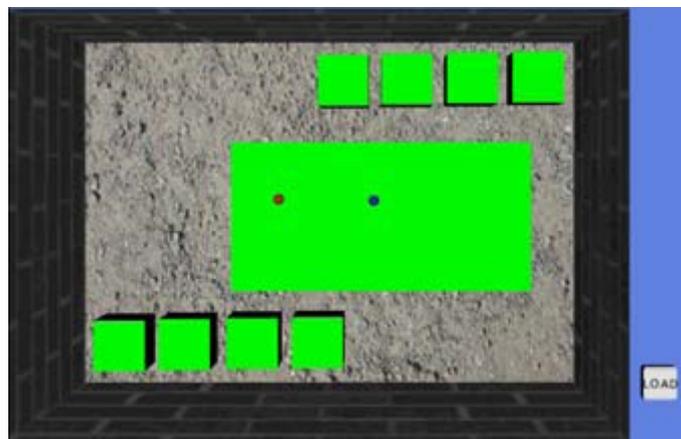
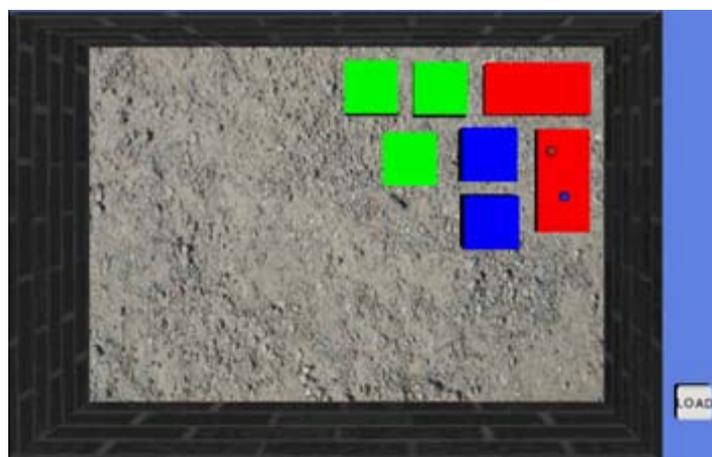


Fig. 2 The table building assignment right after startup.

#### Task 2.

*On the floor there are now seven building blocks of different sizes. Three of these are cubes with volume  $8 \text{ cm}^3$  and the other types of objects have volumes of  $8 \text{ cm}^3$  or  $12 \text{ cm}^3$  respectively. None of the objects can be turned. Your assignment is to build a cube with volume  $64 \text{ cm}^3$  by using all of the building blocks.*

Figure 3 shows the assignment right after it has been loaded.



*Fig. 3 The cube building assignment right after startup.*

## 4. Analysis

A video analysis was performed of the video-recordings from the test sessions in this study. The observation technique was explorative and we were open-minded about which categories to use, because the situation investigated in this study was novel. Thus, the observation of the video recordings of test sessions as well as the analyses of the interview data was not guided by fixed categories. The categories identified as interesting during the analyses however, were the following. Note that the result sections will not be strictly structured according to these categories as certain categories are subcategories of others and also depend on each other in different ways.

- Usability; the effectiveness of the system and the ease of use of the different functions provided
- Common ground; in what ways the collaborating partners was shown to share the same understanding of the task, the layout of the workspace and the objects in it. The use of grounding strategies, such as deictic references, was of outmost interest here as well as the haptic and audio feedback's influence on the dialogue.
- Awareness; in what ways the blindfolded participant understood changes in the interface made by him/herself and the sighted peer. Special focus is placed on how sound and haptic feedback respectively can aid the blindfolded participant in getting an overall picture of what is currently happening in the interface
- Guidance; this is a kind of grounding strategy that is especially important since many of the functions implemented in the application give the ability of providing haptic and auditory guidance, e.g by using deictic referencing.
- Social presence; if the participants feel that they are together in the interface trying to solve a task together. Special focus is again placed on how haptic and audio feedback respectively can create this feeling.
- Modality; how haptic and audio feedback respectively aided the collaboration. This category includes everything that differentiates between the haptic and audio feedback.

The video recordings of the interaction in the shared virtual environment were analyzed. The verbal communication was transcribed and the behaviour was annotated. Annotations were made for each evaluation session describing the users' interaction with the interface and with each other for each group respectively. Everything was done using the video analysis software Transana. Each piece of data was categorized (piece of data was dialogues or description of events). The data, which were first sorted by time, were then sorted into the different categories as a way to ease comparison between the different categories for the different groups and test conditions. Each piece of data was then divided into findings that were unique for audio groups, unique for haptic only groups or general. The data from the different categories were then compared in order to derive general as well as unique findings. In a last iteration we extracted particular dialogue examples as a way to illustrate our findings.

The interviews were transcribed in their entirety. Annotations were made

for each meaningful unit of the data material. The data was analyzed in this way for each group. Finally, the results from all groups were compared in order to obtain general findings and interesting patterns in the results as well as unique but yet informative findings.

## 5. Results from observations of video recordings

In this section we will present and elaborate on the important findings derived from the qualitative analyses of the video recordings from the experiment. We will not only focus on differences between the two conditions visual/haptic/audio and visual/haptic respectively but also on other interesting findings such as grounding strategies and ways of giving guidance. In the dialogue examples, the sighted participant will be denoted *S* and the blindfolded *B*. The observations, analysis and results are thus solely based on qualitative data. Quantitative results from the study are presented and discussed in another article, (<reference to own work>).

The result section will start with a general discussion about different usability aspects and will thus address the usability category. We will next consider common ground, “the shared understanding of the workspace”, and in that part of the result section we will also address the categories guidance and social presence. Last, we address the awareness category under the title “the shared understanding of the workflow”. The shared understanding of the workflow was one of the most important aspects of awareness in our current study. Since the category modality, concerns all other categories the results regarding that category will be covered throughout the result section.

### 5.1 Aspects of usability

Generally, across both experiment conditions, the blindfolded had no problem in using the haptic equipment. Evidently it was also easy to feel things and to distinguish between objects and their textures and heights. The sighted participants had a harder time using their equipment, much due to the PHANTOM Omni having much lower resolution and update frequency than the PHANTOM Desktop, used by the blindfolded participants. However the difference between the equipments gave the blindfolded participant an interesting advantage – he/she could feel details like height differences much easier. Example 1 below, illustrates a typical situation, taken from an audio group. The blindfolded participant is currently on one of the cubes.

Example 1:

**S:** *Can you feel how big they are?*

[Blindfolded moves around the cube for a while]

**B:** *These might be 2x2x2*

**S:** *Yeab, I guess so. but it seems we have five of these 2x2x2 actually, but they are in two different colours, green and blue, so there has to be some difference between them*

- [Sighted feels on both the cube and a tall block, but is not successful]
- B:** *The blue blocks, where are they?*  
 [Blindfolded was guided to one of the tall blocks and moves up and down on it for a while]
- S:** *What is the difference?*
- B:** *The height is more on this one*
- S:** *It's high?*
- B:** *Yeab it's longer*
- S:** *Ok, it's higher. then the blue one should be the 2x2x3.. So what we have is like two of the 2x2x3 and three of the volume of 8, so they are cubic which means 2x2x2 and ... did you get to feel any of the bigger ones?*  
 [They now move on discussing the long blocks and their orientations]

In our setting, described earlier the sighted was watching an ordinary, upstanding, computer screen. This means that the sighted saw the blindfolded's avatar moving up and down on the screen as the blindfolded moved backward and forward on the floor level. Thus they had different perspectives; the sighted participant saw the floor in front of him while the blindfolded felt the floor when she pushed from above. This issue was a bit troubling for some pairs in the beginning but they all adapted to it quite quickly.

The functions we referred to above as the haptical guiding functions were used in most groups, some used them more than others. It was easy and fruitful for most groups to hold on to the same cube. Grasping the other person's proxy was harder however, due to the fact that the forces applied to the respective haptic devices were a little too high. When using that function the devices were often shaking. When the participants managed to use this function however, it was shown to be a great aid. By using haptic guiding either by holding on to the same object or holding on to the other's proxy, users could focus more on the actual task and less on verbal guiding. From the observation analyses it was obvious that all users understood how to use these functions and that they understood, due to the haptic feedback, that they "were in contact" or held the same object.

### 5.1.1 *The audio feedback*

Generally all participants in our audio groups used at least some of the audio cues to communicate with each other or acquire awareness of changes. We can conclude from the observations that they were easily understood and that they could be used in the intended way. We will describe how the participants used these audio cues in a later section.

Although the sound was 3D we noticed that it was hard to hear front/back and up/down. For example, when they used the contact sound the blindfolded could position her/himself correctly in the left/right direction but additional verbal feedback was needed to guide in the other directions. However, this shortcoming did not seem to disturb the collaboration that much and the contact sound was still widely used.

## 5.2 Shared understanding of workspace

Generally, the participants in a group seemed to share an accurate view of the interface with the objects in it, regardless of whether there was sound present or not. The following dialogue should serve as a good example:

Example 2:

- B:** *Can you try like,.. you have one that is 3 high, and then,.. well, one of these 4 long you have that one in the bottom and you put one that is 2 high on top of that and then...*  
[S places a cube (2-high) a tall block (4-long)]
- B:** *...to the side of that you put the one that is...*
- S:** *3*
- B:** *"..3.. and then you put the one that is 4 on top of it*  
[S moves everything into place]  
[They then move on to form the final result]

The participants in this particular group shared a common view about the different objects and how the different dimensions would add up. Actually, the blindfolded was the one solving the assignment. This was actually the case for almost half of the groups, something that says a lot about the accuracy of the shared view of the environment. The participants were often highly engaged in discussions about dimension and how different blocks could go together. As stated earlier, these discussions are mostly based on what the blindfolded feels for the moment.

### 5.2.1 Grounding strategies

The groups used different grounding strategies in attempts to reach a shared understanding of the workspace's layout and the objects in it. These strategies did not seem to be affected by the experiment condition. Almost every group started out, in both assignments, by going through each object haptically while discussing them. Example 1 above is a typical example.

Example 3 also illustrates this:

- S:** *I'm just trying to find out which blocks are the highest...*  
[S feels around and ends up on a high block]
- S:** *... yes, the blue ones are highest*
- B:** *Ok*
- S:** *and the other one..*  
[B moves around and ends up on a long block]
- S:** *there are two... you're at a red one nom..*
- B:** *That one?*

[He points repeatedly to the long block to clarify]

[He then moves along the long and short edges, respectively]

**B:** *quite short.. quite long*

**S:** *Yes, and the height..*

[B moves up on the box and down again a few times]

**B:** *short!*

**S:** *Yes, the low ones... The other one is.. no, up in the corner*

[B navigates to the forward right corner]

**S:** *There it is*

**B:** *Ok*

**S:** *This is the other way..*

**B:** *Yes, turned 90 degrees left,.. ok*

It is interesting to consider how the haptic feedback is utilized during grounding. As can be seen in the above example the blindfolded uses the haptic feedback to point out the object he thinks they are talking about. He does so by touching the object repeatedly from above. The touch modality makes it possible to establish a common frame of reference in a rather straightforward way. Without the haptic feedback this grounding dialogue would probably have been much longer. You can also see from the above example that the sighted talks about what the blindfolded currently feels – the haptic feedback makes it possible to focus on a particular object at hand, something that is very important when grounding. It is also clear, from the above example, that the use of deixis is made possible thanks to the haptic feedback: The sighted refers to “this” a number of times. This also says something about how the haptic feedback influences the feeling of social presence. As the blindfolded feels on the objects the sighted describes them, like they were actually there together wondering around in the room. Last, example two shows how one can refer to boxes. This particular group referred to the objects by their colour, but the most usual way was to use the dimensions like in example 1 above.

A few groups used a different, but yet interesting grounding strategy. The sighted participants arranged the different types of blocks in the second task so they were lying side by side. In this way the blindfolded could easily feel the height relations.

The most common way of maintaining a common frame of reference and a shared view of the interface was to ask questions. The blindfolded repeatedly asked questions about what was currently felt or what had been built so far. Sometimes the sighted also asked clarifying questions like “Do you feel the floor?” The participants often reverted back to discussions about the different dimensions and the blindfolded was, especially in the second task, often guided around what had been built so far.

### 5.2.2 Guiding strategies

The participants in the audio groups could communicate in three ways; verbally, haptically and with audio. The other groups could only communicate verbally and haptically. Verbal communication was by far the most utilized form of communication, especially in the beginning

of the tasks when the participants had to establish a common view of the interface and the overall situation. The haptic feedback was a good complement to the verbal communication as discussed earlier. Audio can also be used as a means of communication, e.g. by using the contact sound to indicate one's own position.

### 5.2.2.1 Verbal guidance

The following dialogue shows a very typical example of verbal guidance. Here the blindfolded participant is guided when attempting to place one of the long blocks in the soon to be cube.

Example 4:

- S:** *Ok, right,.. right, right,.. left,.. feel, .. you can feel the corner*  
[B moves to the upper right corner]
- S:** *... and then go to the left a bit.. stop!.. down.. left, left,.. down,.. go down till you feel the ground.. and then go forward until you feel the wall...*  
[B moves the block to the upper wall, sliding it on the floor]
- S:** *... and right until you feel the boxes,... so, yeah,... let go*

As can be seen in example 4 verbal guidance works, but it can be quite cumbersome. In this group the participants made use of the direction words right, left, up and down (up and down meaning forward and backward in the room respectively). Other groups used points of the compass and some used forward/backward instead of up/down. The important thing here is that every group managed to establish a common frame of reference regarding direction words.

As can also be seen in the above example the participants often refer to interface elements like “the floor”, “the wall”, “the corner” and so on. This guiding strategy was also utilized in most of the groups and of course referring to interface elements makes the collaboration a whole lot easier. Obviously, using this strategy for guiding would not work without the haptic feedback – the sighted can not refer to the interface elements like in the above example if the blindfolded can not feel them.

### 5.2.2.2 Haptic guidance

There were two types of haptical guiding functions introduced to the participants in the training task. One could either hold on to the same cube or grasp the other person's avatar. The former way of giving haptic guidance was the one most widely used. In almost every group the participants placed the board together to complete the first assignment. This was not necessary, of course, but it was a way of involving the blindfolded participant in each step of the solving process. In the second assignment the participants placed at least some blocks together in most of the groups. Again, this was an easy way of involving the blindfolded in the work process. As we saw in example 4 verbal guidance could be very cumbersome. If there were no way of guiding

or communicating haptically solving tasks together would be much harder. In that case the sighted could be tempted to do everything by himself.

The second function, grasping the other's avatar, was used by only a few groups. This was much due to this function being unstable. However, example 5 below highlights the true potential in this kind of guiding function. In the example the sighted participant had just built an L-shape consisting of two long blocks. He had also placed some other blocks close to the L-shape.

Example 5:

- [S grabs B's avatar]  
[He drags B to the beginning of the L-shape]  
**S:** *Non, here we have an L-shape..*  
[S drags B to the top of the shape]  
**S:** *... this is the top.*  
[S now drags B back and forth on the L-shape's  
north-southern part a few times]  
[He then drags B to the east, until the shape ends]  
**S:** *Ok, and this is the bottom right... and then we have this cube that is taller than the others*  
[He drags B up and down on a tall block placed beside the L]  
**S:** *We have another one just like it*

As this example illustrates one can actually show things in a very physical way, by grabbing the other person. This means of communicating haptically conveys a lot of information, which does not need to be spelled out verbally. This way of communicating or guiding also increases the feeling of social presence when using the interface.

Apart from using the methods described above some groups invented their own ways of utilizing the haptic modality for guiding purposes. For example, in one of the groups the sighted pushed the blindfolded's cube with one of his own cubes in the first assignment. This was a way of giving guidance without actually touching the blindfolded's cube directly which was not allowed in this task. Again, the haptic feedback was used to communicate direction information. In another group the sighted participant held her avatar in the "air" to create a physical stop, helping the blindfolded to put the table legs at appropriate places. These examples show how one can use haptic feedback in order to communicate information in these kinds of interfaces.

### 5.2.2.3 Guiding with sound

The contact sound was used regularly as a way of communicating in at least half of our sound groups. Despite the fact that it only worked well in the left/right direction this sound turned out to be a great aid when communicating information. One of the groups actually used it more often than verbal guiding. The sighted participant in this group used verbal guiding as a complement to the sound or when the sound did not work (like in the forward/backward direction). If we look at example 4 again we can see that a big part of the dialogue could probably

be replaced by some uses of the contact sound. Example 6 below shows how the contact sound was used.

Example 6:

- S:** *Pick up a new cube*  
[B locates a cube on her own]
- B:** *That one?*
- S:** *Yeah...And then you can move here...*  
[S uses sound to show the way]  
[B navigates to a place slightly above the intended one]
- S:** *Ok, down a bit..., down..., stop*  
[B releases]
- S:** *You can try to pick up the cube that's here...*  
[S uses the contact sound again]  
[B navigates to the exact place in a few seconds]

As can be seen in example 6 the contact sound conveys a lot of valuable information that otherwise had to be given verbally. As could be concluded from the observations, the blindfolded participants were guided more often in the sound groups. This is probably because the sound cue decreases the work-load on the person who has to guide. Thus, the way of giving guidance and the work-load on the one who has to guide is affected by the addition of a sound cue. That the sound affects the dialogue in a subtle but yet interesting way is also made clear.

It is also interesting to consider the use of deixis, like “come here” and “the cube that’s here”, in the above example. The sound can also be used, with some limitations, to point in the interface and as an aid when discussing different objects. The sighted can use the sound and the blindfolded his avatar as a pointing device. In this respect the sound also makes a positive contribution to the feeling of social presence, since the sound, as well as the haptic feedback, make it easier to focus on and discuss different parts of the interface together

### 5.3 Understanding the workflow

When it comes to awareness aspects the results show that audio cues give a positive contribution. These kinds of cues can inform a person that something is changing and make the blindfolded participant aware of work in progress. Results also show that it is possible for the blindfolded participant to obtain a good understanding of the workflow – the blindfolded knows where the objects are and what to do with them and understands the status of the work.

#### 5.3.1 *The significance of audio feedback*

The haptic feedback can be used to track changes while feeling around on the floor – one can feel if something has been moved or is showing up at a new place. The blindfolded participant then need, however, to constantly explore the whole work-space, something that became evident when studying the groups who did not have access to audio feedback. In the audio groups, however, the sounds instantly inform the blindfolded participants if something has been moved. They do not need to explore actively to find out if something changed if they get audio cues and they are always aware of the fact that work is in progress. Interestingly enough, one of the blindfolded participants in a sound group felt a little anxious when she had not heard anything for a while, and she asked her peer “So, what’s happening now?”. It had been silent for a while because the sighted was just looking at the situation without moving anything. In the groups without audio questions like “Have we started?” and “Are you doing anything?” were frequent, indicating that the blindfolded was unaware of that work was in progress.

Another advantage of audio cues is that they give the blindfolded participants a confirmation of what they are doing themselves. It was clear from our observations that the force feedback generated when lifting an object did not give enough information to the blindfolded participant. The same could be said about the dropping of a cube. In our haptic only groups, questions like “Did I drop it?” and “Did I pick it up?” were frequent. Obviously, the touch feedback was not enough to confirm the actions. These kind of questions were never asked in the audio groups. The sound was enough confirmation of that the blindfolded participants did what they intended to do. It would definitely be possible to do things without audio feedback, but the sense of control would probably be affected in a negative way. Example 7 below shows yet another advantage of the collision sound.

Example 7:

[B accidentally drops the cube she is holding]

[The collision sound is heard]

**B:** *Oh, I dropped it now*

**S:** *Yes, you can pick it again if you feel it*

[B picks up the cube again]

In the above example the collision sound informed the blindfolded about the cube being accidentally dropped. She did not have to wait for the sighted to point this out. In some of the haptic only groups, however, it took quite a while for the sighted to realize that a cube actually did not follow the blindfolded’s avatar. In one case the blindfolded dropped a cube several times in a row without noticing it. Each and every time he had to wait for the sighted to tell him that the cube was dropped.

The difference between touch down sounds did also convey valuable information, despite the fact that they were pretty similar. This is illustrated in example 8, where the blindfolded is placing a cube on another cube to build a leg.

Example 8:

[B puts the cube down, floor collision heard]  
[B puts it down again, floor collision again]  
[B lifts and places the cube on another, object collision heard]

**B:** *Does it look nice?*

**S:** *No...*

**B:** *But I know it's on top...*

In this case the sighted never had to direct the blindfolded for her to be able to place the cube on another one – the sound was enough. In the haptic only groups the sighted had to give verbal guidance while the blindfolded was moving up in the air. It is easy to see that sound makes a big difference for the better in this case. The blindfolded participant could actually do some things completely by his/herself – he/she is not completely dependent on the sighted peer.

Example 9 illustrates another advantage given by the audio feedback. The example shows that sounds give the sighted participant important information as well. He tells the blindfolded to grasp a certain cube, he waits until he hears the grasp sound after which he starts to guide.

Example 9:

[B is feeling around on a long block]

**S:** *Ok, you can pick that one up again*

[B continues feeling the object for about 10 seconds]

[B grasps the object, grasp sound is heard]

**S:** *Ok*

[S starts guiding to the correct place]

In the haptic only groups the sighted participant had to ask questions like “Did you drop it?” or “Did you grasp it?” all the time. Since those kinds of questions were never asked in the audio groups we can conclude that the sound gives enough confirmation and that it gives valuable information to both the blindfolded and the sighted participant. Thus, the sound simplifies the collaboration – the blindfolded participants can concentrate more on the task than on the constant need to ask and in other ways find out what happens.

That sound increases the understanding of the workflow in our application is fairly obvious. It does also increase the work efficiency, providing information that otherwise had to be conveyed through verbal guidance and communication. The sound model gives the blindfolded participant a sense of what the other person is doing and that work is in progress as well as confirmation of his/her own actions.

### 5.3.2 *The benefit of haptic feedback*

It is interesting to consider the way in which the haptic modality influences the understanding of the workflow. If the haptic feedback was not available it would probably be hard to track changes and to do things together. Everything would have to be spelled out verbally by the sighted.

Probably the sighted would prefer to do everything by him/herself, excluding the blindfolded from the work process.

One clear advantage of haptic feedback can be studied in example 3 given above and in example 10 below. The blindfolded can point at and feel objects as the sighted speaks about them making both participants aware of the fact that they are talking about the same thing. Of course, for this to work, both participants have to be aware of the visibility of their respective avatars. In example 10 we can also see how the blindfolded uses his avatar to direct the sighted person's attention to the object at hand. This was a strategy used by the blindfolded participants in every group.

Example 10:

- S:** *Everything that is 3 high has to be combined with 1*
- B:** *Yeab*
- S:** *And there are two of them right now*  
[B moves up to the front right corner]
- B:** *One in the corner here*
- S:** *Yes*
- B:** *And the other one should be here*  
[B moves around on the floor in the backward left corner of the soon to be cube]
- S:** *No*
- B:** *Closer to me.. so that this one can be on top..*  
[B points to a long block more than 1 dm below!]  
[B then moves with his avatar back and forth on the left side of the soon-to-be cube to show what he meant with "on top"]
- S:** *But we have only two of those that are 3 high*
- B:** *Yeab*
- S:** *...but we have two of the flat ones*
- B:** *Yes.. I think I have a pretty good picture of how it should be*
- S:** *Ok, so where do you want to put it? There is something wrong now, one is standing on the floor and the other one on the long one*  
[S refers to the 3-high (tall blocks)]  
[B points to the one standing on the ground]
- B:** *I want to put it closer to me.. It should be close to me and next to the lying one in the corner*
- S:** *Next to the lying one?.. What you currently have is one flat one lying in the corner*
- B:** *Yes, and on top of that there should be one three high and one two high*  
[That is the current situation]
- S:** *Yes*
- B:** *And then, going towards me from the 2 high you should have a 3 high standing on the ground*  
[Now, the sighted sees the solution and the remaining blocks are placed]

This example is also yet another illustration of how the blindfolded actually solves the second task. Note the communicative function of the haptic feedback, especially when the blindfolded

directly points to the intended object lying more than 1 decimetre below. The example also shows that the participants share a common view of the environment and the overall situation.

The haptic feedback could also, to some extent, provide information about the whereabouts of the others. For example, some of the sighted participants made use of the small force generated when the avatars coincide. This was a way of telling the blindfolded that they were both on the same object. The forces generated when lifting the same object also conveys information to both participants that they are actually collaborating on the task. Last, the sighted participant sometimes pushed from the opposite side of a block that the blindfolded held to tell where he/she was or to guide in a certain direction. Apart from conveying information about the other person's whereabouts the haptic feedback also, of course, adds to the feeling of social presence.

## 6. Results from interviews

In this section the result from the post-test interviews will be presented and discussed. See the beginning of section 5 for an explanation of the outline and how it relates to the categories derived from the analysis.

### 6.1 Aspects of usability

Generally, the participants found it fun and interesting to work with the system – a new way of using the computer and “taking the mouse to the next dimension”. Most of the groups also liked the challenge of communicating when one of the co-workers had lost a vital sense. The most valuable function was, according to most participants, that you could actually feel something. When we asked about which type of feedback gave the most valuable information almost everyone answered the haptic feedback. Haptic feedback was identified to be absolutely necessary for the blindfolded and a good complement for the sighted.

When it comes to the sound feedback we got mixed results. Most of the audio groups made use of the audio feedback and the participants understood how to use the different kinds of sound. The grasp and contact sounds were identified as most helpful, but yet problematic. It was particularly hard for the blindfolded to hear where the sound came from in the backward/forward direction. A few blindfolded participants also complained about there being too many sounds present at the same time. The fact that there was no difference between sounds from the two users was also seen as problematic. Here is an example dialogue from the interviews illustrating the sound feedback problem:

**S:** *Well, the audio was kind of hard to use because it was hard to shon,.. I'm here.*

**B:** *Yeah, it was a little bit hard to tell, actually.*

**S:** *...but it helped sometimes to get her to a location.*

Working with different perspectives of the virtual environment was also seen as hard by most participants. They often confused the direction terms in the beginning. One blindfolded participant said the following:

*The thing with these words,.. because you wanted to say "up" and I wanted to say "in front" because the thing I see is this flat thing here in front of me and she sees it upside down on the screen. But I didn't know you saw that so I didn't know why you wanted to say "up".. I didn't think about that until now,.. cause I never saw the screen and I didn't know that..*

The haptic feedback was found easy to use for the blindfolded participants, they could discriminate between objects and textures without problem. Some of the sighted participants had a harder time feeling details, though. This is due to the lower resolution of the PHANTOM Omni.

When we asked the participants about ideas for improvement they came up with different kinds of sound cue refinements in the audio groups. In the haptic groups, however, almost everyone said they wanted sound cues like a contact sound and sounds from colliding surfaces. This is quite interesting since that is just the type of cues we have implemented.

## 6.2 Shared understanding of the workspace

The participants in almost every group felt that they were working in the same environment together, feeling that they were both socially present. This was much due to the blindfolded being able to point to and feel on objects while the sighted talked about them. The blindfolded participants also felt (and in some cases heard) the sighted person's changes in the environment and the participants could also feel each other's proxies when they were close.

In most of the groups the participants did not believe that they had the same view of the environment. This was often explained by them having different perspectives. In most groups, however, the participants experienced that they had a shared understanding of the layout of the room and the objects in it. Most of the groups established this view when the blindfolded was feeling around on objects as the sighted described them. Thus, the haptic feedback was a great aid when the participants needed to establish a common understanding of the workspace and the objects in it.

When it came to the tasks the vast majority of the groups had a common view about the first one. It was seen as pretty simple and they could imagine in their minds what the result should look like. In the second task it was harder. Especially the blindfolded participants did not believe that they had the same idea about how to tackle it. The second task was seen as very complex and it was hard to plan a solution in advance.

Some groups highlighted the importance of having a common ground about direction terms (left, right, etc.). They also found these common terms to be a frame of reference when collaborating. Some groups also mentioned the importance of having the same object references:

- S:** *Yeab, I had to use some way to,.. instead of using the dimensions of the boxes which was quite hard to explain I used colours just to assign a colour with a dimension.*
- B:** *Yeab, I remember you told me "this is the red one" and I could feel it first and I remember the shape.*

### 6.3 Understanding the workflow

In all haptic only groups and in the majority of the audio groups the blindfolded participant reported that they did not have an accurate understanding of what the sighted person was doing. In the sound groups the blindfolded participants could hear that their co-worker was doing something, but they did not have a very good sense of what was done and which object was moved. One respondent reported:

*Even though I could hear that she was somewhere to the right or to the left of me it doesn't help because,.. probably because I'm not blind, this feel for me wasn't really clear where all the cubes were and where everything was so it's not like I had this perfect image. If she is picking something up "Oh, she's probably lifting the cube in the right corner", I couldn't figure that thing out from my picture, so I was concentrated on what I was doing instead of trying to figure out what she was doing.*

In the haptic only groups, the only ways of tracking the other person's actions was to be at the same place or holding the same block. Most of the participants in those groups said that they tracked the other person's actions primarily by talking and not so much by using the haptic feedback.

Most of the blindfolded participants also found it hard to track changes in the environment. Once again the blindfolded participants in the sound groups knew that something was changing but not what it was and where it happened. In the haptic only groups the only way to track changes was to constantly move around feeling what had been moved. One blindfolded participant complained about her constant loss of orientation when objects disappeared and ended up at different places:

*The first time when you go and feel the cubes and so on it's good but when the other one goes and replace it somewhere else its way hard to find or know where it's placed or where should I look? So you lose your coordination in the room where you are, so you must go and seek, "where is the back wall and the corners?"*

Even though it was hard for the blindfolded participants in the audio groups to track changes and his/her partner's actions the sound feedback was identified to be an aid in several cases. Most importantly the sound feedback made the blindfolded participants aware of that work was in progress, something that gave them comfort. They also appreciated the feedback they got on their own actions – they did not have to ask the sighted if they had grasped or dropped an object. A majority of the blindfolded participants in the haptic only groups, however, complained about

not getting enough feedback on their own actions. Another, positive, issue brought up during the interviews was that the contact sound was a great aid, although it could have been easier to use it with ear-phones or surround speakers.

In the haptic only groups the majority of the blindfolded participants believed that sound cues would help them to keep track of the other person's doings. One sighted participant from an audio group reported that the sound helped him to track the blindfolded's actions – he did not have to constantly look at what the blindfolded was doing. We will conclude this section with some very informative words, uttered by a sighted participant in an audio group.

*With no feedback it wouldn't be possible but without sound I think it would be possible but it wouldn't be as easy, and it would be kind of boring for me because I had no idea of what he was doing if I hadn't had the sounds.*

## 7. Discussion

In this article we have described and elaborated on the qualitative results of a study comparing a visual/haptic interface with a visual/haptic/audio one. The results clearly show that the added audio functions make a difference.

One of the biggest problems identified during the evaluation in the earlier study by (<reference to own work>) was the difficulty the visually impaired user experienced when trying to understand the workflow. The most important reason was that objects could be moved around by all users present. Thus the visually impaired user did not have total control. In our haptic only groups we noticed the same problem – the blindfolded participants in these groups had to repeatedly ask what was going on and they seemed more unaware of what was going on. In collaborative settings this is, of course, a major problem. In our sound groups, however, the results were more encouraging. The participants could use the sounds to communicate information and even though the 3D-model could be greatly improved the sounds were obviously helpful when it came to acquiring an understanding of what went on. Even though the blindfolded participants in these groups were not constantly active, the sounds gave them comfort in knowing that work was in progress and that the system was running. Thus, one conclusion from our study is that the presence of sound makes a considerable difference when it comes to the understanding of the workflow.

It has to be pointed out however, that the sound model used is rather simple and that it has its limitations. Sometimes it was obviously hard for the blindfolded participant to hear where the contact sound came from. This is rather problematic since this sound is used for locating the sighted user. However, we could also see the potential of this kind of sound. It was especially hard to know if a sound came from the front or the back. But then again we could clearly see the usefulness of this type of sound. If we improve the sound model and add ear-phones we are convinced that the results would be even clearer, at the benefit of the audio feedback.

We also concluded that audio feedback did not have as much impact on the shared understanding of the workspace's layout as the haptic feedback had. We did show, however, that the sound had an impact on the dialogue. In the haptic only groups much time was spent on asking clarifying questions like “did I drop it now?”. In our audio groups these types of questions never arose, since the blindfolded participants could hear if they e.g. picked up or dropped an object. Thus, the sound affected the dialogue and made it easier to concentrate on the actual task. This probably holds true for other types of interfaces as well that supports collaboration between sighted and visually impaired persons.

The interface could have benefited from more earcons providing details on demand, as was shown to be beneficial in both (Zhao et al., 2008) and (Wall and Brewster, 2006). Specifically, it would be interesting to evaluate functions that could provide a gist (overview of the contents of the interface) of all objects or objects in a specified area in the interface. The microphone metaphor evaluated by Crommentuijn (2006) would also be interesting to evaluate in the context of collaboration between sighted and visually impaired persons, to see if it would ease the grounding process. Last, since it has been shown in a number of studies that haptic feedback is most suited for exploration and navigation and audio feedback more suited for detailed

information (see i.e. (Yu and Brewster, 2003)), the grounding process could be even more simplified if earcons were provided that gave information about the dimensions of a specific object, when the user demands that information.

The haptic feedback also had an important impact on the dialogue – the possibility to feel the objects enabled a more direct discussion. The sighted participant could talk about what the blindfolded felt for the moment and adapt his way of talking to the blindfolded's haptic model. The participants also used numerous deictical references made possible by haptic feedback, which facilitated the collaboration and allowed them to focus more on the task. Even though we could not see that the sound made a difference regarding the shared view of the workspace the results clearly highlights the potential of haptic feedback in that area. These findings correspond well with our earlier study (<reference to own work>), where we tested almost the same interface with groups of sighted and visually impaired children. The haptic feedback's influence on the dialogue and the overall collaboration is considerable and we argue that this holds true even in cases in which groups of sighted persons are collaborating, at least in interfaces where both do not have the same view or where objects are hidden. More studies have to be performed to show the true potential of the communicative aspects of haptic as well as audio feedback. However, the audio functions used in the present study provided mainly what we call “interaction sounds”, which are very subtle. The sounds used do not contain any verbal or clearly symbolic information, making them more unobtrusive than earcons.

When it comes to social presence our results correspond well with earlier research (Oakley et al., 2001; <reference to own work>). The haptic feedback had a positive impact on the feeling of social presence. The participants discussed the objects as they were together in that particular room. Being able to touch each other and feel each other's forces on shared objects they are also given the impression of being co-located. This feeling of being socially co-located ought to be even more important when you cannot see. The haptic feedback makes possible a means of interaction that would otherwise be much more difficult to achieve. We could not see a big difference regarding social presence between the two conditions. This is probably because the sounds are quite artificial and you do not feel like you are co-located with your avatar (you are not surrounded by sounds, like you were actually in the room). This highlights an important aspect of virtual environments – the level of immersion. If the participants wore ear-phones they would feel more immersed in the virtual environment and they would probably think of the sounds as coming from the respective avatar rather than the computer. It would be interesting to conduct new studies with a higher level of immersion to see if the role of the audio feedback changes.

Last, it has to be pointed out again that blindfolded sighted people participated instead of visually impaired ones in our experiment, mainly because we wanted to have a larger data sample to analyse than in our earlier study (<reference to own work>). This may of course have affected the result to some extent. However it also has to be pointed out that most of our findings regarding haptic feedback and its communicative and navigational functions made in (<reference to own work>) correspond well with the results from the present study. When it comes to the added audio cues it is harder to estimate the effect of using blindfolded sighted persons, especially since visually impaired persons may interpret spatial sounds in another way. However, since we were mainly interested in communicative aspects of the sound and how the

sounds could be used for giving feedback on actions, we would not expect that it would be much of a difference if truly visually impaired persons had been participating instead.

## References

- Basdogan, C., Ho, C., Srinivasan, M.A., Slater, M., 2000. An experimental study on the role of touch in shared virtual environments. *ACM Transactions on Computer-Human Interaction* 7(3), 443-460.
- Bergamasco, M., Avizzano, C., Dipetri, G., Barbagli, F., Frisoli, A., 2001. The museum of pure form: system architecture. In: *Proceedings of 10th IEEE International Workshop on Robot and Human Interactive Communication (Bordeaux, Paris, September 18-21, 2001)*. IEEE, New York, pp. 112-117.
- Bergamasco, M., Prisco, G., 1998. Design of an anthropomorphic haptic interface for the human arm, In: Shirai, Y., Hircse, S. (Eds.), *Robotic Research, the Eight International Symposiums*, Springer, London, pp. 278-289.
- Burke, J., Murphy, R., 2007. RSPV: An Investigation of Remote Shared Visual Presence as Common Ground for Human-Robot Teams. In: *Proceedings of the ACM/IEEE conference on Human-robot interaction (Arlington, Virginia, March 8-11, 2007)*. ACM, New York, NY, pp. 161-168.
- Clark, H. H., 1996. *Using Language*. Cambridge University Press, New York.
- Clark, H. H., Brennan, S.E., 1991. Grounding in Communication. In: Resnick, L., Levine, J., Teasley, S. (Eds.), *Perspectives on Socially Shared Cognition*. American Psychological Association, Hyattsville, MD, pp. 127-149.
- Crommentuijn, K., 2006. Designing Auditory Displays to Facilitate Object Localization in virtual haptic 3D environments. In: *Proceedings of 8th ACM Conference on Computers and Assessibility (Portland, Oregon, October 22-25, 2006)*. ACM, New York, pp. 255-256.
- Dourish, P., Bellotti, V., 1992. Awareness and coordination in shared workspaces. In: *Proceedings of the SIGCHI conference on Human factors in computing systems (Toronto, Canada, October 31-November 4, 1992)*. ACM, New York, NY, pp. 541-547.
- Frisoli, A., Simoncini, F., Bergamasco, M., 2002. Mechanical design of a haptic interface for the hand. In: *Proceedings of the 27th Biennial Mechanisms and Robotics Conference (Montreal, Canada, September 29-October 2, 2002)*. ASME Press, New York, pp. 25-32.
- Gaver, W.W., Smith, R.B., O'Shea, T., 1991. Effective sounds in complex systems: The ARKola simulation. In: *Proceedings of SIGCHI conference on Human factors in computing systems: Reaching through technology (New Orleans, Louisiana, April 27-May 2, 1991)*. ACM, New York, NY, pp. 85-90.

- Hoggan, E., Brewster, S., 2007. Designing audio and tactile crossmodal icons for mobile devices. In: Proceedings of the 9th international conference on Multimodal interfaces (Nagoya, Japan, November 12-15, 2007). ACM, New York, NY, pp. 162-169.
- Jansson, G., Juhasz, I., 2007. The reading of virtual maps without vision. In: Proceedings of XXIII International Cartographic Conference (Moscow, Russia, August 4-10, 2007). Available on Conference CD.
- Kennel, A.R., 1996. AUDIOGRAF - A diagram reader for blind people. In: Proceedings of The Second Annual ACM Conference on Assistive Technologies (Vancouver, Canada, April 11-12, 1996) ACM, New York, pp. 51-56.
- Kirk, D., Rodden, T., Fraser, D.S., 2007. Turn it This Way: Grounding Collaborative Action with Remote Gestures. In: Proceedings of the SIGCHI conference on Human factors in computing systems (San Jose, California, April 28-May 3, 2007). ACM, New York, NY, pp. 1039-1048.
- Kraut, R.E., Fish, R.S., Root, R.W., Chalfonte, B.L., 1993. Informal communication in organizations: Form, function and technology. In: Baecker, R.M. (Ed.), Readings in groupware and computer-supported cooperative work: Assisting human-human collaboration. Kaufmann, San Francisco, CA, pp. 287-314.
- Kuber, R., Yu, W., McAllister, G., 2007. Towards developing assistive haptic feedback for visually impaired internet users. In: Proceedings of the Conference on Human Factors in Computing Systems (San Jose, California, April 28-May 3, 2007). ACM, New York, NY, pp. 1525 – 1534.
- Loomis, J.M., Lederman, S.J., 1986. Tactual perception. In: Boff, K.R., Kaufman, L., Thomas, J.P. (Eds.), Handbook of perception and human performance. Wiley/Interscience, New York, NY, pp. 31.31-31.41.
- Malone, T., Krowston, K., 1990. What is coordination theory and how can it help design cooperative work systems? In: Proceedings of the 1990 ACM conference on Computer-supported cooperative work (Los Angeles, California, October 7-10, 1990). ACM, New York, NY, pp. 357-370.
- McCarthy, J., Miles, V., Monk, A., 1991. An experimental study of common ground in text-based communication. In: Proceedings of the SIGCHI conference on Human factors in computing systems: Reaching through technology (New Orleans, Louisiana, April 27-May 2, 1991). ACM, New York, NY, pp. 209-215.
- Mynatt, E., Weber, G., 1994. Nonvisual presentation of graphical user interfaces: Contrasting two approaches. In: Proceedings of the SIGCHI conference on Human factors in

- computing systems: Celebrating interdependence (Boston, Massachusetts, April 24-28, 1994). ACM, New York, NY, pp. 166-172.
- Neale, D.C., Carroll, J.M., Rosson, M.B., 2004. Evaluating Computer-Supported Cooperative Work: Models and Frameworks. In: Proceedings of the International Conference on Computer-Supported Cooperative work (Chicago, Illinois, November 6-10, 2004). ACM, New York, NY, pp. 112-121.
- Oakley, I., Brewster, S., Gray, P., 2001. Can You Feel the Force? An Investigation of Haptic Collaboration in Shared Editors. In: Baber, C., Faint, M., Wall, S., Wing, A.M. (Eds.), Proceedings of Eurohaptics 2001. University of Birmingham, Birmingham, pp. 54-59.
- Pietrzak, T., Martin, B., Pecci, I., Saarinen, R., Raisamo, R., Järvi, J., 2007. The micole architecture: multimodal support for inclusion for visually impaired children. In: Proceedings of the 9th international conference on Multimodal interfaces (Nagoya, Japan, November 12-15, 2007). ACM, New York, NY, pp. 193-200.
- Plimmer, B., Crossan, A., Brewster, S., Blagojevic, R., 2008. Multimodal Collaborative Handwriting Training for Visually-Impaired People. In: Proceedings of the 26th annual SIGCHI conference on Human factors in computing systems (Florence, Italy, April 5-10, 2008). ACM, New York, pp. 393-402.
- Poll, L.H., Eggen, B.H., 1996. Non-visual Interaction with GUI Objects. In: Sasse, M.A., Cunningham, R.J., Winder, R.L. (Eds.), Proceedings of HCI on People and Computers XI, Springer, London, pp. 159-168.
- Rasmus-Gröhn, K., Magnusson, C., Efring, H., 2006. User Evaluations of a Virtual Haptic-Audio Line Drawing Prototype. In: McGookin, D., Brewster, S. (Eds.), Proceedings of the 1st international workshop HAID 2006. Glasgow, UK, August 31 - September 1, 2006, pp. 81-91.
- Saarinen, R., Järvi, J., Raisamo, R., Salo, J., 2005. Agent-based architecture for implementing multimodal learning environments for visually impaired children. In: Proceedings of the 7th international conference on Multimodal interfaces (Toronto, Italy, October 4-6, 2005). ACM, New York, NY, pp. 309-316.
- Sjöström, C., 2001. Designing Haptic Computer Interfaces for Blind People. In: Proceedings of 6th international symposium on Signal processing and its applications (Kuala Lumpur, Malaysia, August 13-16, 2001). IEEE, New York, pp. 68-71.
- Wall, S., Brewster, S., 2006. Feeling what you hear: tactile feedback for navigation of audio graphs. In: Proceedings of the SIGCHI conference on Human Factors in computing systems (Montréal, Québec, Canada, April 22-27, 2006). ACM, New York, NY, pp. 1123-1132.

- Winberg, F., Bowers, J., 2004. Assembling the senses: Towards the design of cooperative interfaces for visually impaired users. In: Proceedings of the 2004 ACM conference on Computer supported cooperative work (Chicago, Illinois, November 6-10, 2004). ACM, New York, pp. 332-341.
- Winberg, F., Hellström, S-O., 2001. Qualitative aspects of auditory direct manipulation. In: Proceedings of the 7th international conference on Auditory display (Helsinki, Finland, July 29-August 1, 2001). ACM, New York, pp. 16-20.
- Yu, W., Brewster, S., 2003. Evaluation of Multimodal Graphs for Blind People. *Journal of Universal Access in the Information Society* 2(3), 105-124.
- Zhao, H., Plaisant, C., Shneiderman, B., Lazar, J., 2008. Data Sonification for Users with Visual Impairment: A Case Study with Georeferenced Data. *ACM Transactions on Computer-Human Interaction* 15(1), 1-28.