Auditory Feedback in Haptic Collaborative Interfaces

YING YING HUANG, JONAS MOLL, EVA-LOTTA SALLNÄS, YNGVE SUNDBLAD
HCI Dept., School of Computer Science and Communication, Royal Institute of Technology, Sweden

The combined effect of haptic and auditory feedback in shared interfaces, on the cooperation between visually impaired and sighted persons is under-investigated. A central challenge for cooperating group members lies in obtaining a common understanding of the elements of the workspace and maintaining awareness of the other members’, as well as one’s own, actions during the work process. The aim of the experimental study presented here, was to investigate if adding audio cues in a haptic and visual interface would make collaboration between a sighted and a blindfolded person more efficient and if it would improve perceived awareness, common ground, and task performance. Results showed that task performance was significantly faster in the audio, haptic and visual feedback condition compared to the haptic and visual feedback condition. One special interest was also to study how participants utilize the auditory and haptic force feedback in order to obtain a common understanding of the workspace and to maintain an awareness of the group members’ actions. Results from a qualitative analysis showed that the auditory and haptic feedback was used in a number of important ways in the participants’ grounding process and for the group members’ action awareness.

Categories and Subject Descriptors:
H.5.2 [Information Interfaces and Presentation]: User Interfaces-Auditory (non-speech) feedback, Haptic I/O. Evaluation/methodology; H.5.3 [Information Interfaces and Presentation]: Group and Organization Interfaces-Evaluation/methodology; Synchronous interaction General Terms: Experimentation, Human Factors, Measurement, Performance

Additional Key Words and Phrases:
Haptic, Force feedback, Virtual environments, Multimodal interface, Collaboration, awareness, common ground

Corresponding author: Ying Ying Huang
E-mail address: yingying@csc.kth.se
1 INTRODUCTION

During computer-supported cooperation, a central challenge lies in increasing peoples’ attention to actions done by those involved in joint work in a shared interface. It is particularly interesting to investigate how haptic and audio feedback in combination support communication and interaction between people during collaboration in shared interfaces in a situation where one person cannot see. Here, haptic feedback refers to an integration of both kinesthetic sensing (i.e. of the position and movement of joints and limbs) and tactile sensing (i.e. through the skin) [Loomis and Lederman 1986]. Auditory feedback refers to the use of non-verbal sound to convey information to users in a computer interface. A haptic device that provides force feedback can assist users in being aware of and to identify objects in a computer interface. In the study presented here, a Phantom desktop and a Phantom Omni were used for generating force feedback with high resolution in three dimensions. Both devices are operated with a pen-like stylus attached to a robotic arm that generates force feedback [Patomäki et al. 2004]. Some information conveyed by haptic feedback, such as location of objects and the location and action of your partner in a collaborative environment, could also be represented by auditory feedback. We argue that auditory feedback might be efficient to further increase people’s awareness of others’ actions in a haptic collaborative virtual environment.

In this paper we present an experimental study in which it is investigated if auditory feedback makes co-located collaboration more efficient in a shared haptic virtual environment. In both conditions (with or without auditory feedback), participants communicated verbally and could touch and manipulate objects in a visual interface with a haptic device. In one of the two conditions, participants could use auditory feedback that gave information about their actions in the virtual collaborative environment. The aim of the experiment was to investigate the effectiveness of adding audio in a haptic collaborative environment, which could be in terms of improving task performance, and increasing perceived awareness, perceived common ground and perceived task performance.
The work presented here was carried out within the EU-funded project MICOLE (Multimodal Collaboration Environment for Inclusion of Visually Impaired Children) in which applications were developed that support collaboration among visually impaired and sighted children [Sallnäs et al. 2007]. The aim of the MICOLE project was to build a multimodal collaborative environment for better inclusion of visually impaired children in group work in different subjects in primary school. Computer supported group work is a much neglected perspective in assistive research and development.

2 THEORETICAL BACKGROUND

2.1 Haptic Force Feedback

Haptic technologies have become widely available and inexpensive, and the advantage of being able to use the touch modality is being recognized. The research concerning haptic perception and rendering techniques has increased rapidly during the last few years, and results have shown the significant role that haptic feedback plays in graphical single user interfaces [Gupta et al. 1997; Hasser et al. 1998; Hurmuzlu et al. 1998]. Users with severe visual impairment have to work without the visual modality, which restrains their ability to utilize graphical user interfaces. Specifically, it is harder using the sense of touch to get an overview of an environment and localize and explore objects and interesting parts thereof [Jansson 2007]. Designers are beginning to realize the advantages of haptic displays in order to help blind individuals to overcome the challenges experienced when accessing and exploring the web [Kuber 2007]. The touch modality has also been shown to make it possible for visually impaired to explore and navigate virtual environments [Sjostrom 2001]. The interaction is enriched by the use of the sense of touch, since visually impaired users can identify objects and perceive their shape and texture. In a recent EU project, Pure Form, the aim was to develop a haptic display for exploration of virtual copies of statues at museums in order to make them accessible to visually impaired people, [Bergamasco et al. 2001; Bergamasco and Prisco 1998; Frisoli et al. 2002].

In a number of studies on collaboration, it has been shown that haptic feedback improves task performance and increases perceived presence and the subjective sense of togetherness for different application areas in shared virtual environments [Ho et al.
Less attention has been paid to the impact of auditory feedback in combination with haptic feedback in a collaborative setting. However, in one study the interaction between visually impaired pupils and their teachers was addressed [Plimmer et al. 2008]. This study investigated the effects of training handwriting using haptic and audio output to realize a teacher’s pen input to the pupil.

2.2 Auditory Feedback

Audio displays are also getting increased attention for their capability to support navigation and collaboration in virtual environments. A number of researchers have investigated approaches with varying degree of success, illustrating how functionality in 3D virtual environments can be made accessible to the sighted and the visually impaired people by providing auditory feedback. Winberg and Hellström [2001] developed an interface based solely on auditory feedback. They developed a sound model that made it possible for blind users to play the game “Towers of Hanoi”. The game was set with either three or four disks and each disk had its unique sound differing in pitch and timbre. The height on the peg of a particular disk was represented by the length of the sound. Stereo panning was used to convey information about which peg a particular disk was on. In that study, the results showed the potential of the audio modality to convey information to visually impaired users, who could play the game together with a sighted person in a way that included both players in the process of solving the problem.

In an experiment conducted by Poll and Eggen [1996], a blind subject used an absolute mouse to scan for graphical user interface (GUI) objects represented by speech and non-speech sounds within a rectangular area bounded by standing edges. Kennel [1996] suggested that blind users, within a relatively short time could read simple diagrams with the aid of a touch panel and an associated auditory display. The diagram was displayed on a sheet of paper covering a tablet. When touched, parts of it generated relevant audio messages. The whole diagram could thus be explored using this audio-tactile strategy.
Visually impaired people get information mostly by hearing and touching and they rely on non-visual media for navigation and collaboration in interfaces. According to the design principles of Winberg [2004], a non-visual interface should provide access to all functionality. In principle all the things that can be done by a sighted user in a GUI should be possible to do with an interface for a blind user (this will include giving a sense to the ‘iconicity’ of the interface and its spatial arrangement). Furthermore, it should enable manipulation and exploration. If possible the direct manipulation methods of GUIs should be given a parallel implementation. Finally, it should be coherent with visual interfaces in order to enable collaboration between sighted and visually impaired users. This is likely to involve the cross-media translation of significant interface elements for creating a common mapping of the elements used in both kinds of interfaces.

An approach of transforming visual information into non-visual media has been used for mapping GUI objects in a study by Crommentuijn [2006]. Five different ways of representing a set of objects for visually impaired users in a GUI were implemented in an auditory and haptic interface and visually impaired users’ interaction with this interface was investigated. 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 one. From the results in these studies it can be implied that information such as the location of objects and the location and action of your partner in a collaborative visual or haptic context could also be represented and conveyed by auditory cues.

2.3 Non-Visual Collaboration

In Swedish schools pupils often do group work in many different topics. This pedagogical approach trains social skills and supports collaborative learning. Assistance in this group work process is particularly important if one of the pupils is visually impaired since the most important sense – vision – is not available. There are several studies which have investigated issues surrounding collaboration between visually impaired people in educational settings. David McGookin and Stephen Brewster investigated an initial study of computer supported collaboration between visually impaired users based around the interactive
browsing and manipulation of simple graphs. They specifically looked at supporting awareness of others' activities and interaction between participants [David McGookin, Stephen Brewster, 2007]. The teaching and learning of subjects in small groups, largely without direct teacher involvement, has been argued to improve the social, academic and cognitive abilities of students [Slavin, R.E. and Cooper, R., 1999]. For visually impaired students, however, there are significant problems in accomplishing these aims. In interviews with visually impaired students, Salinä [5] describes how students enjoyed group work, noting that for one it allowed her to “get to know the other pupils better that she had not talked to so much before”. Salinä also noted however that it was difficult for students to keep track of other’s activities. For example, if two students were browsing a raised paper diagram they would each need a copy, making it difficult for one to point out something to his/her partner in the same way a sighted user might. This makes it difficult for each person to know what his/her partner is looking at, or to be sure both are referring to the same part of the diagram, therefore impairing collaboration.

2.4 Awareness in Dynamic 3D Environments

In real life we can have an extensive overview of the environment we are in and of objects in it. However, we always focus on some objects which attract our attention. Moreover, we attend to objects by multiple senses. In virtual reality environments, multisensory experiences are created through artificial means and the effectiveness of virtual environments has often been linked to the richness of sensory information and realness of the experience [Basdogan et al. 2000; Held and Durlach 1992; Ellis 1992; Barfield and Furness 1995]. The extent to which users are aware of things that happen in virtual environments could also be used to determine the effectiveness of the sensory information.

Awareness is generally used in terms of individuals’ perception of others’ activities and the status of others’ work processes [Dourish and Bellotti 1992]. A key factor to consider in collaborative virtual environments is the way in which continuous awareness of others’ activities allows people to manage their own activity in social situations in a flexible way and to predict actions of others [Benford et al. 1994]. When people who cooperate do not have the
opportunity to get sufficient awareness information, they do not reach the same quality in joint projects [Kraut et al. 1993].

In order to clarify terminology on awareness, we refer to Carroll et al. [2003], who distinguish between three kinds of awareness in virtual settings and emphasize that each type can be supported by certain tools. Social awareness refers to the users’ consciousness of the presence of others. Social awareness can be fostered by tools which visualize the presence of others in any way, for example by using photographs of the team members. Tools fostering action awareness provide information about the actions which are currently being carried out by the group and its members, for example by showing which shared resources (e.g. a document) other team members are interacting with. Accordingly, action awareness is primarily important in the context of synchronous collaboration. Activity awareness focuses on the task that is to be performed by the group. Here, the actions of the participants involved are related to the mutual task, but activity awareness is primarily important in the context of asynchronous collaboration. Carroll et al. [2003] presented this theory of awareness in the context of computer-mediated communication in a distributed environment.

Many studies have paid attention to awareness and its impact in distributed virtual reality systems [Benford et al. 1994; Kraut 1993; Kimmerle 2007] and in the context of computer supported collaborative work (CSCW) [Schmidt 2002; Neale et al. 2004; Carroll et al. 2006].

We believe that the concept of action awareness is also relevant in a co-located situation, especially when one person cannot see. When users are working on a joint task in a shared virtual environment, such as for example putting together a model of a machine in a shared interface, people need not only to communicate verbally and to see each other and the environment with the model. They also need to feel haptic and hear auditory feedback in order to convey their own intentions, to understand others’ intentions, and to coordinate joint actions. In a co-located situation, which was investigated in our present study, action awareness of the things that group participants do in the physical as well as the virtual environment has to be attended to by all participants in order to be able to perform the task.

In collaborative situations, sighted people mostly rely on visual information whereas visually impaired people mainly attain
awareness by touching and hearing. It is necessary to transfer visual information, at least to some extent, into non-visual representations in order to make computer-supported communication and collaboration between sighted and visually impaired people possible. In this case getting action awareness depends on multimodal technical support. Haptic and auditory feedback systems give sensory information that potentially can provide crucial cues for visually impaired users that make mutual awareness during collaborative work possible.

Several questions on users’ action awareness were considered in the course of the present study. How do sighted and non-sighted participants attend to the other and get attention in a joint task during collaborative work? How do the non-sighted manage to keep track of the location of his or her partner’s proxies in a virtual environment? How can non-sighted people use haptic feedback to track changes that are being made in a dynamic environment? Would adding audio, apart from haptic feedback, make a difference in collaborative virtual environments?

2.5 Shared Understanding of the Work Space

Several questions regarding users’ grounding strategies during collaboration in a shared multimodal environment in a situation where one participant cannot use the visual sense were also addressed in this study. For instance, what grounding strategies do participants develop when they use different sensory modalities in their collaborative work? Could adding audio feedback, apart from verbal, serve as an information source that participants could use as one more resource in the grounding process?

Clark and Brennan [1991] define common ground as a state of mutual understanding among conversational participants about the topic at hand. People must have this shared awareness in order to carry out any form of joint activities.

Although common ground is a general theory of language use, it holds true for all collaborative activities that people must update their common ground on a continuous basis, and do so through a grounding process [Neale et al. 2004]. To communicate, collaborate, and coordinate, people must share a vast amount of information and mutual knowledge. We argue that group members can achieve common ground by a variety of approaches, not only conversation
but also by different sensory information like touch and interaction sounds.

Neale et al. [2004] suggested the term “activity awareness”, incorporating the term activity from the very broad and muti-layered concept activity theory. In order to more fully understand the role activity awareness plays in remote collaboration, he developed a model for evaluating activity awareness (Figure 1). In this model the most important variables that need to be considered when evaluating activity awareness are presented. It focuses on the central relationships underlying the processes of distributed collaboration. We argue that this model can be applied to a co-located group work situation as well. Neale’s model illustrates the important factors needed for understanding the relationships between variables that are important for collaboration. The key preconditions that the model points out is that if the proper levels of communication and coordination are supported, groups can achieve common ground and acquire activity awareness efficiently.

Fig.1. Model for evaluating activity awareness with the factors needed for understanding the relationships between variables that are important for collaboration by Neale et al. [2004].

In the current study, the central questions are how multiple sensory cues with haptic and auditory grounding tools can improve mutual understanding in order to enrich communication and enhance
coordination, and how they potentially make collaboration more efficient in a situation where one of the persons cannot see the workspace. To our knowledge, the impact of audio feedback on efficiency and satisfaction in haptic collaborative interfaces has not been investigated in detail so far.

3 AIM AND HYPOTHESES

In our research we want to investigate whether adding audio functions in a visual and haptic interface will make a difference beyond the impact of haptic feedback. A high level aim was to investigate the impact of auditory feedback on awareness of actions made in a virtual environment and on the grounding process. The hypothesis tested in this experiment was if adding audio functions to a collaborative visual/haptic interface would improve task performance in a collaborative haptic 3D virtual environment:

(H1) Adding audio feedback into the collaborative visual/haptic environment will make task performance faster.
(H2) Adding audio feedback into the collaborative visual/haptic environment will increase users’ perceived awareness.
(H3) Adding audio feedback into the collaborative visual/haptic environment will increase users’ perceived common ground.
(H4) Adding audio feedback into the collaborative visual/haptic environment will increase users’ perceived task performance.

4 METHOD

4.1 Experimental Design

A between subjects design was used in this experiment, with two conditions: (1) a visual and haptic VR environment, and (2) an audio, visual and haptic VR environment. The dependent variable was task performance. Task performance was measured by the time spent by group members to solve a task during the test. According to McLeod [1996], the usability of a system can be measured by how long it takes to perform a task and how well the task is performed. It has been shown that the larger the number of modalities, the higher performance would be expected from the participants [Short et al. 1976].

The test sessions ended with an open form of interview with each pair. Questions were asked in the interview about the subjects’
perception of the system, with special focus on awareness, common ground, and joint task performance in different modalities. An observation analysis of the video recordings was also made in order to get a more detailed understanding of how audio cues affected the interaction.

4.4 Subjects

There were 32 students participating in this experiment and divided into 16 pairs. One subject was sighted and one subject was blindfolded in each pair when the test task was performed. Eight of the pairs used the interface with audio, visual and haptic force feedback and the other eight pairs used the interface with only visual and haptic force feedback. The participants were matched in such a way that they knew each other well since before. The argument was that it would be better for them to collaborate with a familiar person, whom they had known for a long time and felt comfortable working with. Visually impaired people were not recruited to this experiment even though it would have been better than blindfolding sighted people. More participants were however needed, than could be recruited among visually impaired. In basic research regarding the effects of auditory information on the time to perform two joint tasks, it can reasonably be assumed that the effects are the same on visually impaired and blindfolded sighted people. The general level may be different, but if a parameter has an effect on non-handicapped people, it can be expected to also have an effect on visually impaired people.

4.5 The Application

The collaborative interface used in this experiment was modeled to be perceived as a room viewed from above through a transparent ceiling. The collaborative interface and the setting of the experiment are shown in Figure 2. The room contained cubes that could be picked up and moved around by means of touch feedback using a haptic feedback device called Phantom. The roof, walls and the cubes all have different textures that can be felt. The small and differently colored spheres shown in Figure 2 represent two users holding the same object. In this way the users can co-operate in
compiling larger objects. Since gravity and mass is applied to all objects, the users feel the weight and inertia of objects as they carry them around. Besides feeling and manipulating the cubes, users can feel and also grasp each other’s graphical representations in order to provide navigational guidance e.g. to a blindfolded partner. The users can also “feel the other proxy” by means of a small vibration, applied whenever the users’ graphical representations get close enough.

In the visual, haptic and audio interface, a number of auditory functions were added that gave different kinds of audio cues. A “grip sound” was heard every time an object was lifted. In this way one participant could know when the other person lifted an object. The second auditory function was a “touch down sound” which was heard every time an object fell on the floor. In order to distinguish between the sound an object makes when it falls to the floor and the sound an object makes when it falls on another object, a “collision sound” was also designed which was heard every time an object landed on top of another one. The forth auditory function was a “contact sound” which was heard every time the button on the phantom was pushed down except, of course, when an object was grasped. This “contact sound”, in stereo, was heard from one’s own avatar and made it possible for the other participant, and especially the blindfolded one, to know the location of the other user’s position relative to his/her own position.
Fig. 2. The experimental setting with participants using one Phantom Desktop and one Phantom Omni respectively and a picture of the screen with spheres representing participants in the shared environment.

4.6 Apparatus
The hardware used in this experiment were one personal computer with two dual core processors, a computer screen, a computer mouse, a keyboard, and a pair of loud speakers. Two different haptic devices were used; one Phantom Desktop (Figure 3) and one Phantom Omni (Figure 4). Reachin API 4.1 and Microsoft Visual Studio 2003 .NET were used as software platform and CamStudio was used for screen capturing in order to record the interaction in the interface during the experiment.

Fig. 3. One Phantom Desktop force feedback device (source: SensAble Technologies Inc).
4.7 Procedure

The researcher gave introductory information about the aim of the experiment, followed by an instruction on how to use the haptic devices. The researchers made sure that both participants were fully aware of how the haptic and audio feedback worked and could be utilized in the interface before the participants started solving the tasks.

The experiment was divided into four parts: one demo session, one training session, one group work session, and one interview session. In the demo session the soon to be blindfolded participant got the chance to use the Phantom Desktop while looking at the screen. In the demo environment he/she could feel several boxes with different textures and surfaces applied to them. Both participants then practiced on a training task in the experimental environment before the real task, in order to get used to this type of interface. 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. It was made sure that the blindfolded participant felt comfortable with the blindfold and got used to working in this kind of haptic environment before the real task was loaded. After the training session, each pair of
participants solved a task in either the visual/haptic or the visual/haptic/audio condition. The visually impaired participant used a Phantom Desktop and the sighted used the Phantom Omni device in all sessions. Interviews were conducted after the task was performed in order to acquire information about system usability and subjective perception. In the interview open questions were asked to each participant of the pair.

4.8 Task

In the task (Figure 5) the participants were to build a table. The virtual environment contained eight cubes for the four legs and a flat board. Collaboration was inevitable since they were only allowed to move four of the cubes each. However, they were allowed to collaborate in lifting the board. The two participants solved the tasks collaboratively.

“In this assignment eight cubes of size 2 cm x 2 cm 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 cm x 6 cm. 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 precisely placed at the respective corners of the table. 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 move your own cubes.”
Fig. 5. Two users represented by one proxy each in the shape of spheres, lifting an object together during the table building assignment directly after startup.

4.9 Quantitative Analysis

One-way ANOVA (Analysis Of Variance) was used for analyzing the variable time. The data regarding time to perform tasks from one group in the audio/haptic condition was excluded due to technical problems with the application. Therefore, the data a group in the haptic/visual condition was also excluded.

The post-test interviews were transcribed and a qualitative analysis was performed.

An observation analysis was also performed of the video recorded material from the experimental sessions. The dimensions of focus in the qualitative analysis were awareness and the participants' grounding process.

4.10 Qualitative analysis

4.10.1 Awareness

The definition of awareness in this study was “individuals' perception of others' activities and the status of the joint work-process”. In the interview and the observation analysis a number of aspects were considered. How the participants could maintain a shared focus of events in the joint task. How they perceived where the other person was located. How they could track movements of the other person continuously and know the status of ones' own actions.

4.10.2 Common Ground

In this study common ground was defined as “a state of mutual understanding among conversational participants about the topic at hand” [Clark and Brennan, 1991]. In the interview and the observation analysis common ground was addressed using a number of dimensions. Participants shared understanding of the layout their shared understanding of the objects’ qualities. The participants' grounding strategies used different sensory modalities and the extent to which sensory feedback helped users to have the same understanding about the shared context.
4.10.3 Task Performance

In the interviews and in the observation analysis task performance a number of aspects were considered. How well they understood the system, and to what degree they felt that they learned how to use the system, as well as their skill level in using specific features in the system.

5 RESULTS

5.1 Task Performance

The hypothesis was that adding audio functions in the visual/haptic environment would improve task performance. The result showed that task performance, defined as the time to complete the task, differed significantly (p<0.05) across the two conditions. The mean value of the task completion time was shorter in audio/haptic/visual condition (M=581 seconds, Sd= 284) than in the haptic/visual condition (M=916, Sd= 277), (Table 1). This means that subjects used about 10 minutes to perform the task in the audio/haptic/visual condition and about 15 minutes in the condition with no audio feedback.

<table>
<thead>
<tr>
<th>Task Performance (sec.)</th>
<th>Audio/Haptic Feedback Mean (Sd) n</th>
<th>Haptic Feedback Mean (Sd) n</th>
</tr>
</thead>
<tbody>
<tr>
<td>F=5.0 p=0.045*</td>
<td>581 (284) 14</td>
<td>916 (277) 14</td>
</tr>
</tbody>
</table>

Table 1. Experimental results regarding total time to complete task for 14 groups.

*=significant at 95% level

5.2 Result from Observations and Interviews on Joint Action in a Shared Workspace

When going through the video recordings special attention was paid to participants' notion of awareness, common ground and task performance. One aspect of awareness is how the blindfolded participant tracked changes in the environment and how well he/she knew the partner’s whereabouts and doings. The focus was also on how and with what modalities the sighted participant communicated
information to the blindfolded participant in the two experimental conditions when performing joint tasks. We have written another paper focused especially on video qualitative analysis for this experiment [Jonas, al, et. 2009].

Here, a couple of dialog examples have been chosen for this article in order to illustrate the findings regarding participants’ perceived awareness, common ground and task performance as related to the impacts of different multimodal interfaces.

Besides using the haptic feedback cues to pick up and move objects, the sighted participant often used the contact sound to show where a certain cube could be found and placed respectively. Below is a typical conversation when a group tried to solve a task on building cubes. In the example, the potential of the added audio feedback for navigation in a haptic interface is made clear.

Sighted:
Pick up a new cube!
[The sighted used contact sound to express where a cube was and the Blindfolded located it on her thanks to the sound]
Blindfolded:
That one?
[The Blindfolded picked up the cube, and a grip sound occurred]
Sighted: Yeah...And then you can move here...
[Sighted uses sound to show the way again]
Blindfolded navigates to a place slightly above the intended one
Sighted:
Ok, down a bit..., down..., and stop.
[The sighted tries to help the Blindfolded to navigate verbally]
Sighted:
Ok, drop it.
[The Blindfolded releases the cube on another one, and hears a collision sound]
Sighted:
You can try to pick up the cube that’s here...
[Sighted uses the contact sound again]
Blindfolded navigates to the exact place in a few seconds

This is an example of how the sighted user manages to get the blindfolded person’s attention and guide the blindfolded to a certain place. We can also see from the example that the blindfolded got to the right place very quickly. In the groups without audio, the sighted participant often had to spend a lot of time explaining in words (like
“up”, “down”, “too much, go back”) in what direction to move and the blindfolded had to go back and forth numerous times before ending up at the intended place. Furthermore, the grip and contact sounds made the blindfolded participants aware of what the sighted did and approximately where he/she was. All participants in our audio groups preferred to work in a virtual environment with audio feedback, reflecting their actions, rather than in silence. This was also highlighted in the interviews with the sighted and blindfolded participants in the audio groups.

The blindfolded participants in the haptic-only groups, however, complained about not knowing what was going on. They felt that they were working in silence and that they could not contribute in the work process. If sound is not available the only way for the blindfolded to be aware of what is going on is by talking to the sighted peer, who cannot tell every detail about what is happening. The blindfolded participants in the haptic-only groups also did not hear when they had done something by themselves – they had to ask their sighted peers if they had dropped an object, since the touch feedback apparently was not enough. Based on this discussion and the above example we argue that the audio cues make a considerable positive difference regarding the amount of awareness information accessible for the blindfolded and sighted participants in the collaborative work.

Another result from the observations was that our audio groups could use the audio feedback as an aid when establishing a common ground. In most of our audio groups the sighted participant used speech and the contact sound in combination to describe the workspace’s layout and the objects in it, as the dialogue example below shows.

Sighted:
Come here, there are three green cubes in the right up corner.
[Sighted uses contact sound to guide and it works]
[Blindfolded explores the cubes by touch for a while]
Sighted: There are two of the same
  Height there, that one,..., it looks like they have the same height.
Blindfolded:
Yeah, it feels so to, they are the same.
[Blindfolded feels around on the cubes for a while]
After a while the sighted directs the blindfolded to the rest of the building blocks. In the meantime the blindfolded tried to feel the shapes and orientations of the blocks. After a few minutes of exploration the blindfolded picks up one of the tall blocks.

Blindfolded: Should I put it in the corner?
Sighted: Yeah, the corner seems like a very good idea.
Blindfolded: I take that corner!
[Blindfolded puts down the block in the corner, which she had felt earlier]
[Blindfolded moves around in order to update her mental picture]
Sighted: Can you feel in the corner how it feels, and then we can start to think out how to build this thing.
Blindfolded: At this corner?
[Blindfolded points with her avatar]
Sighted: Yeah
Blindfolded: Is this?
Sighted: It’s the long one
Blindfolded: Oh,...., yeah it sure is
[Sighted shows with contact sound where the other tall block is]
[Blindfolded locates it after a short while]

The above situation is representative for almost all the audio groups – the blindfolded is guided through the different building blocks by the sighted peer. This is done by a combination of speech and the contact sound and the fact that the objects can be explored by touch is of course very important too. As can be seen in the above example the blindfolded can follow the sound quite easily (at least in the left/right direction). Due to this the sighted does not need to verbally direct the blindfolded, but can instead focus on describing the different objects’ shapes.

In most of our haptic/visual groups the sighted had to devote a lot of time on guiding the blindfolded peer verbally, something that was often cumbersome as discussed above. It is interesting to study the participants’ grounding processes that are used to achieve the
same view of the layout and the same understanding of the subtask at hand.

In the interviews with the participants in the audio groups most of the sighted ones highlighted the importance of being able to achieve a common ground by means of audio and haptic feedback in the task solving process. They also appreciated the sound feedback when it came to making the blindfolded aware of what was happening. The blindfolded could use the audio feedback as an aid when navigating and orientating and they felt that it was possible to feel things with the haptic feedback equipments. Interestingly, the fact that the blindfolded participants were unable to see the cubes on the screen, made it possible for them to feel the depth of cubes quite easily, while the sighted that used to disregard the sense of touch before making a judgment, had a hard time estimating the heights. As long as the participants established a common ground by talking, navigating and exploring the work space, they could reach the same level of understanding of the layout, the context and task at hand. The blindfolded could contribute a lot to the final solution, and they sometimes were the ones that solved it by suggesting the correct spatial relation between the cubes.

Many of the participants in our haptic-only groups said that the addition of sound feedback would be a great aid in their collaboration.

The findings in this study showed the importance of the audio feedback when it came to supporting collaboration, although the haptic feedback was an important prerequisite. That is consistent with results from several studies that have shown the importance of the haptic modality in collaborative environments. Our study is no exception, the haptic made it possible for the blindfolded to feel objects and in this way both participants could focus attention on and talk about the same object as if they really perceived that they were working in the same work space. In the dialogue examples given in this section we see utterances like “should I put it in the corner” and “it’s the long one”, showing that the participants shared a common view about the properties of the objects in the virtual environment. The haptic feedback is making this possible.

When comparing the two experimental conditions, we identified a positive qualitative difference caused by the addition of sound cues. In the audio and haptic condition, the blindfolded solved the tasks and navigated not only using their sense of touch combined with the
verbal direction from the sighted participant, but also their sense of hearing. The audio modality made it possible for them to orientate (with a contact sound), to hear approximately what their partner was doing and where he/she was and to hear what they did themselves. The fact that the sighted person did use the contact sound and the blindfolded did interpret this sound cue as being the sighted person’s position in the virtual environment and could move in the right direction, implies that this sound might be an example of a feature that shortened the time a blindfolded person spent navigating which in turn made task performance faster.

In the haptic/visual groups, some blindfolded participants said that they had totally no idea whether they did something (such as dropping an object on another one or dropping an object on the floor) or not. Moreover, they did not know what their partner did. Almost all participants in the haptic/visual groups believed that it would make a positive difference if audio was added in the environment. The blindfolded participants in these groups wanted to know that something was going on, and to get confirmations on actions taken.

6  DISCUSSION

In this paper we have shown examples of how audio and haptic feedback can be utilized by groups in order to collaborate in a CVE by using audio and haptic cues to support awareness, common ground and task performance, in a synchronous co-located context. The data from this experiment showed that the interface, that provided audio as well as haptic feedback, made the groups perform the task significantly faster. The result supports the hypothesis that collaboration would take less time when audio cues provided information about the actions that the two participants made in the haptic virtual environment.

The qualitative results based on the analysis of video recordings and interviews, supports the quantitative findings that auditory functions play an important role for collaboration in multimodal environments. The added audio information improved the possibility to get the attention of the other person and provided an additional aid in the grounding process when participants solved a task collaboratively.

In view of the principles of design for communication and collaboration between sighted and blindfolded people [Winberg and Bowers 2004], the results suggest that the auditory/visual/haptic
interface of our prototype is functional. All the basic actions that could be done by sighted users of the GUI could be done by the blindfolded as well. The auditory and haptic feedback enabled manipulation and exploration in the shared virtual space and it was coherent with the visual interface, which enabled collaboration between sighted and blindfolded.

The results from the questionnaires on the subjective variables regarding perceived performance, awareness and common ground, did not show any significant differences. The use of questionnaires to capture people’s experiences of the quite abstract concepts awareness and common ground is a rather uncharted approach. Those concepts spring from paradigms that are mainly based on qualitative analysis of behavior. The reason why no significant differences were found might be that the questionnaires are either immature or that the approach is not appropriate.

However, an alternative explanation could be that adding audio feedback in a haptic/visual interface made the situation very complex, especially in combination with the fact that users were co-located. It is well known that hearing someone talk makes a big difference for people’s notion of social presence [Short et al. 1976], and it was possible to communicate verbally in both conditions. So, even if the participants used the audio feedback to get the attention of the other partner, they maybe did not consciously separate this from the verbal communication but instead integrated the different auditory cues into one and based their reports in the questionnaire on that. Additionally, the different audio cues could be perceived as confusing. If there is, for instance, too much audio information, as could be the case when both participants drops an object at about the same time and in addition communicate verbally the sounds would overlap and could become confusing.

In this study, the sounds such as the “grip sound”, that was heard every time an object was lifted, the “touch down sound”, that was heard every time an object touched the floor and the “collision sound”, that was heard every time an object landed on top of another object, were designed in a very simplistic fashion. These kinds of sounds were designed only as event oriented cues of actions taken. They did not give any advanced information about the work process. An example of an alternative approach that might be pursued in future studies is using continuous friction sounds, heard whenever a user touches and manipulates an object or surface.
The findings in this study are important to consider when designing multimodal interfaces for collaboration between sighted and visually impaired people in order to make collaborative situations more accessible and possible for contribution by visually impaired people.

ACKNOWLEDGMENTS
This study is part of the EU-funded project MICOLE. Many thanks go to Professor Gunnar Jansson for his warmth in discussions and suggestions and to the participants in the experiments.

REFERENCES


anthropomorphic haptic interface for the human arm, In Y. Shirai and S. Hircse (Eds.), Robotic Research, the Eight International Symposia on Microrobotics and Nanorobotics, 278-289. London: Springer.


David McGookin, Stephen Brewster, 2007 An Initial Investigation into Non-Visual Computer Supported Collaboration, Proceeding of CHI 2007, April 28-May 3, 2007 • San Jose, CA, USA


JANSSON, G. 2007. Haptics as a substitute for vision, in M. A. Hersh and M. A. Johnson (Eds.), Assistive technology for
vision-impaired and blind, London: Springer.


Jonas Moll, Ying ying Huang, Eva-Lotta Sallnäs, Yngve Sundblad, Audio makes a difference in haptic collaborative virtual environments, submitted to Interacting with Computers, 2009.


