How To NOT Hit A Virtual Wall – Aural Spatial Awareness for Collision Avoidance in Virtual Environments

Christian Afonso
University of Hamburg
Vogt-Kölln-Str. 30
22527 Hamburg, Germany
2goncalv@informatik.uni-hamburg.de

Steffi Beckhaus
University of Hamburg
Vogt-Kölln-Str. 30
22527 Hamburg, Germany
steffi.beckhaus@uni-hamburg.de

ABSTRACT

Compared to graphics, sound is still an underused modality for conveying information and providing users with more than just general ambience or targeted sound effects. *Collision notification* is one case of direct aural feedback: The moment a user hits a wall, they hear an appropriate sound e.g. a thump. We tried to go further by using contextual spatial sound to provide *collision avoidance* feedback, which plays continuously in the background, but, unlike ambient soundscapes, reacts accurately and in real-time to upcoming collision hazards. In a first experimental design, we provided directional spatial sound feedback for collision avoidance in a prototypical labyrinth environment and examined the performance and reactions of a group of test subjects, who navigated through the labyrinth. Our initial design already received positive reactions from the subjects and analysis of the performance data shows first results indicating the viability of this kind of spatial sound feedback.

Categories and Subject Descriptors

H.5.2 [**User Interfaces**]: Auditory (non-speech) feedback

General Terms

Design, Experimentation, Human Factors, Performance

Keywords

Collision Avoidance, Collision Response, Collision Feedback, Spatial Sound, Auditory I/O, Context-Aware Computing, Multi-modal interfaces, Virtual Environments, Virtual Reality.

1. INTRODUCTION

Collision notification (CN) is used in several kinds of virtual environments (VEs), ranging from games to scientific simulations. In cases where this feedback is more than just a visual stop of movement when colliding with the surroundings, a user receives additional feedback, such as a loud thump sound (as in the study by Blom and Beckhaus [3]) or haptic force-feedback, depending on available equipment and the implementation. The experience of, for example, an unpleasant collision response might subsequently motivate the user to avoid further collisions.

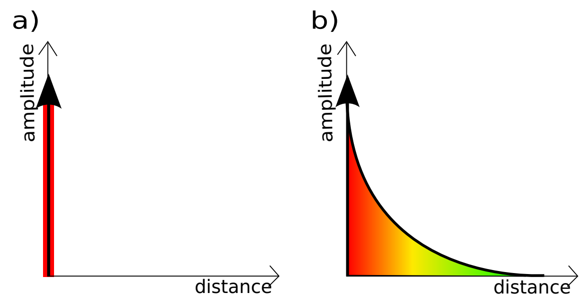

Figure 1: The concepts of collision notification (a) vs. collision avoidance (b).

However, until the moment of collision the user will not be alerted to the imminent danger of impact (Figure 1a). In environments where he or she is moving fast through narrow spaces and is distracted by other tasks such as wayfinding, accidental collision with walls is a constant immediate possibility. Occasional sudden collisions disrupt the flow of interaction, which is especially distracting in applications where a kind of "flow" is desired.

In the real world, we normally avoid objects intuitively. While the main modality for obstacle perception is vision, other senses also play a role, foremost the auditory sense. Hearing, for example, is an excellent developed sense in terms of accuracy, pattern recognition, and speed, compared to the recently most prominently used sense of vision. Studies have shown that blind subjects heavily rely on auditory perception of their surroundings and can detect obstacles in their path with reasonable confidence. They also show that blindfolded sighted subjects quickly begin to pick up these skills [11]. An interesting observation is that the blind subjects often were not consciously aware of using hearing, sometimes describing a general feeling of "pressure" [11, 13]. These findings suggest that all humans are able to perceive and use subtle auditory cues about their surroundings and that these cues are mostly perceived subconsicously.

These kind of cues – changing acoustic characteristics, like the increase in early reflections near walls – currently have no counterpart in virtual environments, although they could provide a subconscious navigation aid. Despite becoming ever more realistic, VEs often fail to give the user a full sense of presence in the environment and the feeling that the environment is real. While physically correct rendering of acoustic reflections in real-time is currently unfeasible, some simplified kind of subliminal spatial audio feedback would reinsert some degree of multimodal awareness of a user's surroundings, as they could hear them in addition to only seeing them.

Therefore, we propose to shift from collision notification alone to tools for *collision avoidance* (CA). The idea is to, instead of letting users only know about collisions with the environment that have already occurred, subconsciously guide them in such a way that they intuitively do not collide in the first place. This is done by providing them with constant feedback of varying, information encoding intensity (Figure 1b). A key point is the use of *spatial sound* for spatial information: the CA feedback should be coming from the directions of the potential collision hazards, providing directional cues that the user hears and can act upon. The sounds used as CA feedback should have such a character that a user is instinctively deterred. We expected an increased navigational confidence due to the implicit subtle notification about a suitable distance, thus a certainty that no surprising sudden collisions will occur.

In the following sections we will give an overview on related research that inspired our work, describe the design of the prototype VR application we built for evaluating the effects of CA feedback, and present the user study we designed to test the prototype application, along with its results.

2. RELATED WORK

Past work on the uses of spatial sound for navigation in virtual environments has mostly focused on providing spatial orientation points, or beacons, used to guide a user along a predefined path or towards specific aural landmarks, or realistically rendering the distribution of sound in an artificial environment. Gonot et al. [5] compare contextualized and decontextualized beacons. Walker and Lindsay [14] studied how different kinds of sounds, different capture radius,

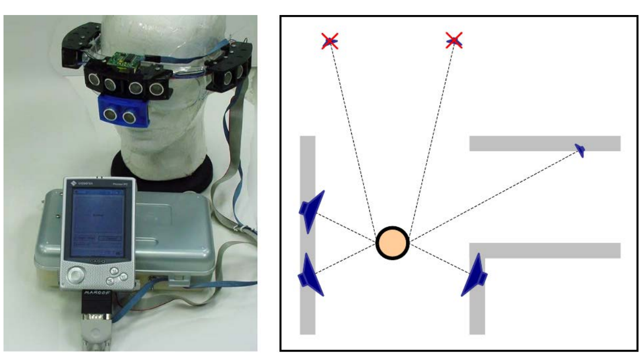

Figure 2: The FIU obstacle detection system. Left: The device used for distance measuring. Right: Sketch of the system's working principle. ***(from [2])***

and practise affected the adherence of test users to a path connecting localized beacon sounds. The AudioGPS implemented by Holland et al. [7] uses a similar kind of artificial directional spatial audio as a real-world orientation and navigation aid. These studies show that users can perceive the direction of artificial spatial sounds well and use them for orientation and navigation.

A recent study by Blom and Beckhaus [3] on the effectiveness of various kinds of collision feedback shows that realistic collision notification feedback helps in making users more alert to avoid collisions with their surroundings. This study was the starting point for our current investigations in collision avoidance techniques.

In a study to assess the fear value of sound parameters for horror games, Garner et al. [4] found that "3D positioning (particularly sound coming from a sharp left or right), pitch (particularly high pitched sound) and loudness (specifically greater relative loudness) to be notably effective in increasing participants' perceived intensity ratings."

Apart from the motivation to extend existing CN techniques, we drew inspiration from real-world applications designed to help blind persons aurally perceive their surroundings. One of these systems is the vOICe [9], which consists of glasses with integrated stereo headphones and a small camera between the eyes; The image from the camera is transformed into a stereo audio signal running from left to right and mapping light areas of the image to frequency-modulated sine waves. This technique, however, only reacts to lightness, not distance, as depth-detection from a single image is difficult. Ifukube et al. [8] describe the use of ultrasound emitters and receivers modeled after the echolocation of bats, combined with a downsampling of the signals to human hearing range, to aid in the detection and location of obstacles in the user's path. Another interesting early system called *Navbelt* was developed by Borenstein et al. [10]. It consists of a belt fitted with eight sonar units at the front side, directed at evenly spaced angles. The distance data from the eight sensors was transformed into a quick succession of sounds with different amplitudes, corresponding to the distance measurements of the sonars, and played back from eight virtual spatialized directions through headphones.

Aguerrevere et al. [2] built a head-mounted rig containing stereo headphones and sonar range sensors, which pointed in six directions around the head, with a portable Pocket PC used as the processing unit (see Figure 2a). The distance measurements from the sonars were used by the Pocket PC to choose fitting spatialized sounds to be played over the headphones, and to control their amplitude. The latter was used to convey the proximity to obstacles (Figure 2b). To spatialize the sounds, a range of sound files with pre-calculated head-related transfer functions (HRTFs) were used, one for each sensor direction. Using these, it was possible to supply the user with a 3-dimensional soundscape, which he could use for forming a mental map of his surroundings.

For an example of environment-aware application of spatial sound, Gonot et al. [5] discuss the *Eye* game, developed by a group of students from the Graduate School of Games (EN-

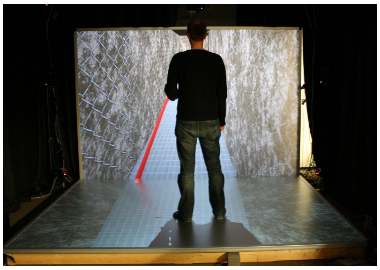

Figure 3: A user standing in the virtual environment projected on the L-shape.

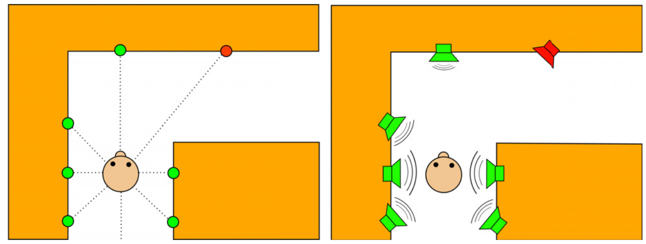

Figure 4: Left: distance sensors. Right: sound source positions.

JMIN)[1]. In this game, the player has to move through a mental asylum mostly with his (virtual) eyes closed, as he can see into the minds of the other inmates and has to protect himself from their traumatic visions. Mostly he moves around relying on his aural sense only, thus the game world is largely communicated aurally. The player has to build a sense of the space he is in by listening to the sounds surrounding him, signifying points of interest as well as danger to keep away from.

3. DESIGN

To examine the viability of the proposed sound supported method for spatial CA feedback in VEs, we implemented an application to operate in our immersive projection-based virtual environment.

3.1 Software, hardware and navigation setup

The application is built on OpenSceneGraph[2] for graphics rendering, VR Juggler[3] for handling input and output devices, and the ACTIF interaction framework [6]. The ACTIF framework provides an abstraction for the steps of interaction processing, ensuring modularity and easy interchangeability of modules, for example for different kinds of input and output devices. We use a Nintendo Wiimote[4] for virtual travel control, and the ARTtrack system[5] for Wiimote and body motion tracking.

The graphics of the virtual environment were projected as stereoscopic images (viewable through shutter glasses) onto the wall and floor screens of our "L-shape" environment (depicted in Figure 3). Four loudspeakers are positioned at the four corners of the L-shape for spatial audio rendering, and low-frequency speakers are built into the floor of the L-shape for haptic feedback. Motion tracking for close-range movement and obstacle dodging is achieved through an optical tracking target mounted on the shutter glasses. The Wiimote is also fitted with a target. Pointing of the Wiimote in a direction and pressing buttons to move and rotate is used for navigating through the environment.

3.2 CA algorithm

The basic functionality of our application is as follows: the geometry around the user's position inside the virtual environment is constantly monitored using distance sensors, implemented via ray picking in 8 directions (as depicted in Figure 4, left – the backwards-facing sensor is not seen in the image, as the wall in that direction is too far away). The number of picking directions (and correspondingly, sound sources) can be set arbitrarily, but during preliminary testing, a number of 8 has shown to be the best tradeoff between directional accuracy and computation load, as the calculations have to be done for each frame.

After the picking is performed, 8 sound sources (one for each picking direction) are updated with the detected positions of ray-wall intersections (see Figure 4, right). The amplitude of a sound source is determined by the distance of the corresponding wall intersection to the user's position. The amplitude is computed by the formula

$$amplitude = (1 - distance)^2$$

where *distance* is normalized to the range [0..1], and 1 corresponds to a customizable maximal distance. If the detected wall distance is above the maximal sensing distance, the corresponding sound source is turned off (indicated in red). The effect is an amplitude distribution as pictured in Figure 1b, with the maximal amplitude directly on collision, and 0 at and above the maximal picking distance.

3.3 The feedback mode

The main operating mode of the CA application uses a single sound file played from 8 different directions, as described in the previous section. This feedback mode was the central subject of evaluation of the prototype, which we tested in this user study.

4. USER STUDY

In our user study we wanted to evaluate the effectiveness of the sound-based mode of CA feedback. For this, we chose three different sounds that we expected users to perceive as unpleasant or unnerving, and that would thus lead users to try and avoid them on an instinctive level. We used a labyrinth-like close-spaced virtual environment (also used in the collision notification study by Blom and Beckhaus [3]) for users to travel from one end to the other while trying to avoid colliding with the walls. Different labyrinth layouts were used for each condition, with all of them having equal length, with a single path leading from start to goal without any forks in the path (The user path plot in Figure 5 shows a typical labyrinth layout).

[1]http://www.enjmin.fr
[2]http://www.openscenegraph.org/projects/osg
[3]http://www.vrjuggler.org
[4]http://www.nintendo.com/wii/console/controllers
[5]http://www.ar-tracking.de

4.1 Selection of sounds

For the CA feedback in the user study we chose three sounds: a constant electric buzz, a deep synthetic bubbling, and a melancholic organ-like minor chord. These sounds were chosen after conducting a study with 19 test subjects in which we examined the deterrance effect of various sounds. For this, we measured, how often test users chose them or avoided them, when having to decide between experiencing the one or the other [1]. This study was done in a VE, in which users were automatically moved forward through a straight, corridor-like environment with obstacles in the middle of the path. They had to avoid this obstacle by physically moving out of its way. Two sounds, randomly chosen from a pool of eight sounds, were placed on either side of each obstacle, so that users had to decide each time to avoid one sound and approach the other. Each sound appeared 14 times, paired twice to each of the other sounds (including a silent sound, for isolated reference).

We measured the avoid/pass ratios for each sound through automated logging and collected the users' own perceptions and evaluations of the sounds with questionnaires. The three sounds we chose for the CA study were passed the least often and, at the same time, were rated as relatively uninteresting, matching our desire to choose sounds that would work well in the background without drawing attention themselves.

4.2 Hypotheses

We postulated the following hypotheses about the outcome of the user study, reflecting our expectations about the effect of collision avoidance feedback on users in the mazes:

- **H1: Spatial collision avoidance feedback reduces collisions.** We expect that users will collide with the walls less often when spatial collision avoidance feedback is given than without it.
- **H2: Spatial collision avoidance feedback helps staying away from the walls.** We expect that continuous spatial feedback will be a constant guidance to the users and subliminally motivate them to stay in the middle of the corridor, minimizing the noise they hear from the sounds.
- **H3: The more unpleasant the presented sound is, the better CA works.** We expect that a sound that is perceived as more unpleasant than another will yield better results, both regarding the number of collisions and the average distance to the walls.

4.3 Test conditions and procedure

In the three CA conditions (one for each sound: “buzz”, “bubble”, and “chord”), users were presented with the full omnidirectional sound feedback, as well as a shock response (acoustic and haptic) upon wall collision. As a reference, we also included a condition without CA feedback, with only a “thump” sound (and haptic floor thud) as CN.

Before starting the study, each participant had the opportunity to move through a labyrinth without any collision feedback (notification or avoidance) to familiarize themselves with the environment and controls. Once they felt confident enough in moving through the environment, they had to pass through four consecutive labyrinths, first with the “thump” condition (CN feedback only), then with the three CA conditions (“buzz”, “bubble”, and “chord”, CN+CA feedback), with the task of traversing the labyrinth swiftly until they arrived at the goal. The participants were not told about the concrete matters evaluated (number of collisions, average wall distance, completion time), nor about the purpose of the sounds they were to hear.

4.4 Collected questionnaire data

After each condition, we asked the test subjects to answer some questions (regarding their perception of their environment, the sound they were hearing as collision avoidance feedback, and how they perceived its effect on their navigation through the maze), by giving a rating on a scale between two extremes:

1. ”How realistic did your environment feel?”
   (not realistic at all – very realistic)
2. ”How did you perceive the spatial sound?”
   (very unpleasant – very pleasant)
   (very calming – very stimulating)
   (very deterring – very attractive)
3. ”Did you feel that the sound helped you navigating the maze?” (not at all – very much)
4. ”Were you afraid of suddenly colliding with a wall?”
   (never – very often)

For the ratings, we chose a *semantic differential scale* (see Tullis [12], section 6.2.2), an interval-based numerical scale from 1 to 7. On this scale, 1 marks one extreme (e.g. ”very unpleasant“) and 7 the other extreme (e.g. ”very pleasant“), while 4 marks the neutral point.

At the beginning of the session, an introductory questionnaire was filled out by the participants. It was used to collect general statistical data (age, sex, occupation, gaming experience, experience with 3-D and VR environments, musical experience) that could be used to further analyze possible demographical effects on the test results.

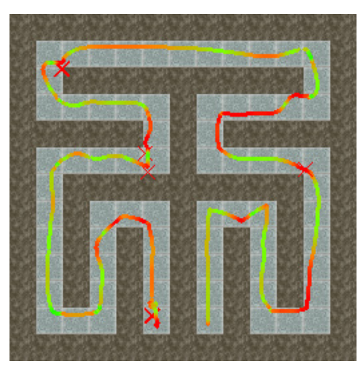

Figure 5: Example of a path plot image.

4.5 Automated performance logging

During each test run we recorded live log data of the user's movement, both as raw data in a text file and as a path plot image. The path plot images were intended to visualize the traveled path for easier manual evaluation and as a simple visual reference. An example path plot can be seen in Figure 5: The path the user took is displayed in colors from a continuum of green (optimal wall distance, middle of corridor) to red (very close to the walls). The color of the path corresponds directly to the maximal amplitude of the sound heard by the test user: where the path is green the sound was very low, where the path is red the sound was very loud (The directionality of the sound is not displayed – the sound was always loudest from the nearest wall). Collisions are marked with a red "X".

5. RESULTS

We had 19 test subjects participate in the user study over the course of three weeks of testing. Most of the participants were students in their twenties, with the average age being 24. More than half of the participants were recruited for testing from a youth orchestra, the rest were mostly students of informatics or related sciences. Slightly less than half (42%, 8 participants) were female, 11 were male.

We asked the participants, how important they found the role of sound in games (as the most widely available and familiar form of interactive applications), for supporting the atmosphere as well as for enhancing gameplay. On a 7-point semantic differential scale between "unimportant" and "very important" the users rated the importance of sound for atmosphere 5.95 on average, while the average rating for the importance of sound for gameplay was 5.37. According to a paired-samples t-test this difference is highly significant (p=0.0041).

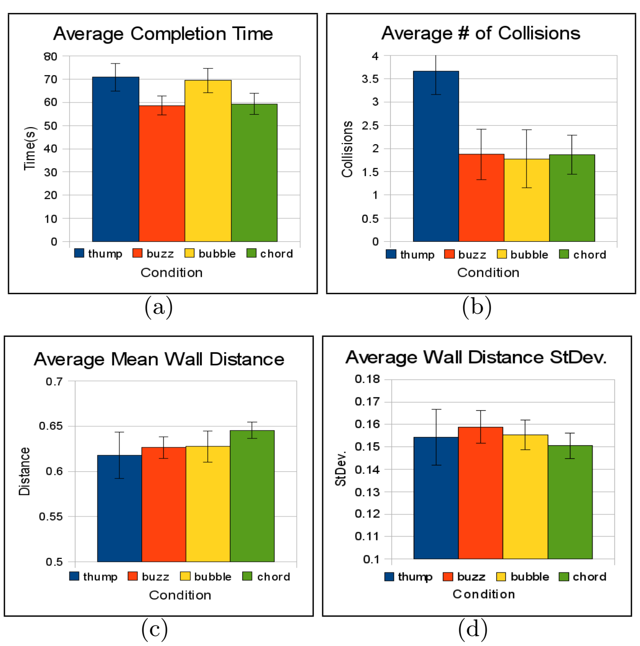

Figure 6: Evaluation results from log files.

5.1 Logging results

After postprocessing the individual log files for each user, we had four complete sets of condensed results, corresponding to the four conditions: the means and standard deviations of wall distance, as well as completion times and the number of collisions. From this user-specific data we got our sample averages, as seen in Table 1.

Table 1: Sample averages of log results.

			Wall Distance	
Condition	Time	Collisions	Mean	StDev.
0:thump	70.83	3.67	0.62	0.15
1:buzz	58.63	1.88	0.63	0.16
2:bubble	69.44	1.78	0.63	0.16
3:chord	59.27	1.87	0.65	0.15

The results are visualized in Figure 6. As can be seen in Figure 6b, the average number of collisions when spatial collision avoidance feedback is given (conditions "buzz", "bubble", and "chord") is only half of when only collision notification is provided ("thump"). Average completion time (Figure 6a) varies between 60 and 70 seconds, depending on the condition, with the "buzz" and "chord" conditions generally resulting in faster completion than "thump" and "bubble".

The average of the per-user mean wall distance (depicted in Figure 6c) increases slightly from the first condition ("thump") to the last ("chord"). The range of possible wall distance was from 0 (collision) to 0.75m, as the corridor was 1.5m wide. The average standard deviation of wall distance (Figure 6d) rises minimally from the "thump" condition to "buzz", the first presented collision avoidance condition, and then decreases a small amount again from "buzz" over "bubble" to "chord".

To see if there were significant performance differences between the conditions, we performed paired-samples t-tests between each pair of conditions for each set of evaluation results. The results of the t-test can be seen in Table 2. Here, **bold** entries mark a highly significant difference (>99% confidence, p<0.01), *italicized* entries mark significant difference (>95% confidence, p<0.05).

Table 2: Paired-sample t-test results of log results.

			Wall Distance	
Conditions	Time	Collisions	Mean	StDev.
Thump-buzz	**0.0028**	**0.0013**	0.5955	0.5042
Thump-bubble	0.8382	*0.0125*	0.7931	0.9819
Thump-chord	0.0596	**0.0030**	**0.0004**	**0.0002**
Buzz-bubble	**0.0020**	0.6282	0.3110	0.2931
Buzz-chord	0.7396	0.8395	0.0744	0.0562
Bubble-chord	**0.0001**	0.6173	0.8666	0.8400

The differences in completion time are highly significant between "thump" and "buzz" (p=0.0028), "buzz" and "bubble" (p=0.002), and "bubble" and "chord" conditions (p=0.0001). Regarding the number of collisions, the "thump" condition is highly significantly different to "buzz" (p=0.0013) and "chord" (p=0.0030), and significantly different to "bubble" (p=0.0125).

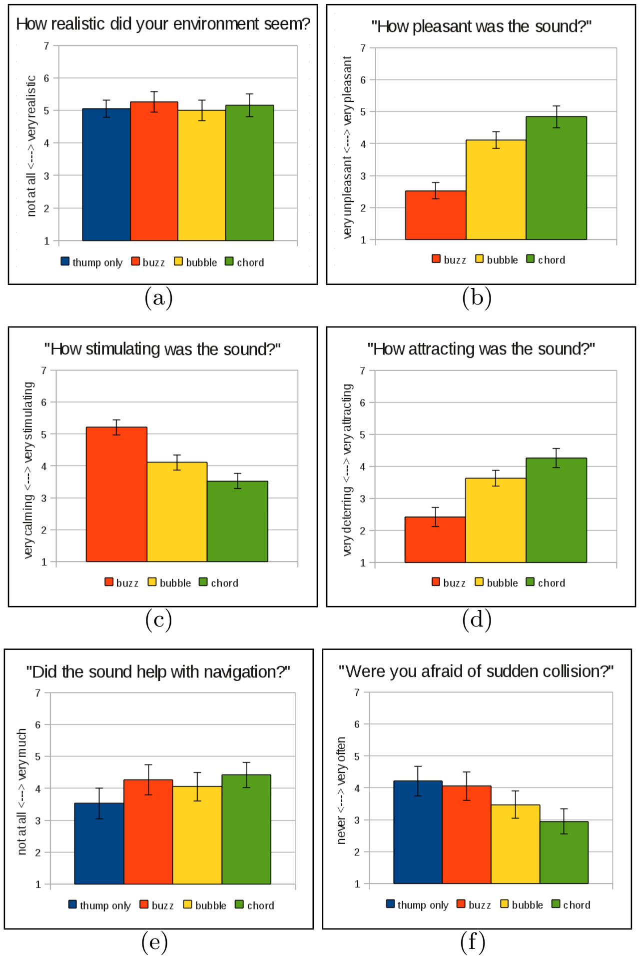

Figure 7: Results from the maze test questionnaires.

In the t-test results of mean wall distance, between the "thump" and "chord" conditions a (highly) significant difference is found (p=0.0004). The difference in average wall distance standard deviation is also highly significant (p=0.0002).

5.2 Questionnaire results

The results from the questionnaires answered after each condition are shown in Table 3 and visualized in Figure 7. All answers are on a semantic differential scale from 1 to 7; the extreme values are next to the vertical axes in the figures.

As with the log data, we applied a paired-samples t-test between each pair of conditions. During all conditions, the participants found the environment to be reasonably realistic, with ratings close to 5 (Figure 7a). According to the t-test result, the difference between the "buzz" and "bubble" conditions, however small, is slightly significant (p = 0.0207), with the "buzz" condition yielding a slightly higher perception of realistic surroundings.

Table 3: Average ratings from maze questionnaires.

condition	real.	pleas.	stim.	attr.	nav. help	coll. fear
thump	5.05				3.53	4.21
buzz	5.26	2.53	5.21	2.42	4.26	4.05
bubble	5	4.11	4.11	3.63	4.05	3.47
chord	5.16	4.84	3.53	4.26	4.42	2.95

As is visible in Figure 7b, compared to "bubble" and "chord" the "buzz" sound was found relatively unpleasant, with a rating of 2.53 compared to 4.11 (bubble) and 4.84 (chord), both close to the neutral rating of 4. The differences are highly significant, with p-values of 0.0002 (bubble) and 0.000003 (chord). As seen in Figure 7c, the "buzz" sound was also rated relatively more stimulating (at 5.21), compared to both "bubble" (4.11) and "chord" (3.53); again, the differences are highly significant with >99% confidence (p=0.0015 between "buzz" and "bubble", and p=0.000022 between "buzz" and "chord"). The difference between "bubble" and "chord" is also slightly significant, with a p-value of 0.0447.

Also, the "buzz" sound was found to be more deterring (with a rating of 2.42, see Figure 7d, with "bubble" being only slightly deterring (3.63) and "chord" being slightly attracting (4.26). The differences are again very significant, with the t-test between "buzz" and "bubble" yielding p=0.0050, and the test between "buzz" and "chord" yielding p=0.000008.

Figure 7e shows the ratings for how much the test subjects perceived the sounds (collision notification in the first condition, and the combination of collision avoidance and collision notification feedback in condition 2 to 4) to aid them with navigation. The "thump" condition, providing collision notification (CN) only, was generally rated slightly lower than the other three, at 3.53. The CA conditions are all around an average rating of 4, halfway between the "did not help at all" and "helped very much" points, with ratings of 4.26 (buzz), 4.05 (bubble), and 4.42 (chord); however, these slightly better rating (compared to "thump") did not result in a significant difference according to the t-test.

Finally, fear of sudden collision decreased over the course of the conditions (Figure 7f), with "thump" yielding a moderate collision fear rating of 4.21, while "buzz" was rated 4.05, "bubble" was rated 3.47, and "chord" was rated 2.95. With the "thump" and "buzz" ratings being very close together, both are significantly different to "chord", the lowest rated sound, with p-values of 0.0212 (thump-chord) and 0.0197 (buzz-chord).

6. DISCUSSION

Although some participants mentioned a slight dizziness after spending some time in the VR environment, only two of them stated that they had become nauseated. Most of the participants adapted quickly to the interaction methods for navigating the mazes. Different persons did, however, show a wide spectrum of behavior for putting them to use, ranging from standing absolutely still and navigating solely with the Wiimote, to moving around quickly on the L-shape platform for short-range movement, dodging walls, and correcting the course they steered with the Wiimote.

On the question of how important sound is for different purposes in games, the participants gave a significantly higher rating for the importance of sound for the atmosphere, compared to its importance for the actual gameplay. However, both ratings are comparatively high in relation to the scale from "unimportant" to "very important". This indicates that, while users are aware that sound plays a significant role in the overall experience, its "active" use as a tool for influencing or aiding gameplay is less appreciated than its "passive" use for setting a mood and establishing the ambiance of a player's environment. Here lies a great potential for enriching the user experience in novel ways.

Regardless of condition, the realism rating given by participants to the environment stayed roughly the same, as visible in Figure 7a. At close to 5 on the 7-point scale, these can be interpreted as the participants perceiving the environment as moderately realistic, being closer to the "very realistic" end of the scale than to the "not realistic at all" end, while the continuous spatial sound feedback in its current form did not affect the perception of realism one way or the other.

6.1 Collision prevention

When looking at the navigation performance of the users in the different conditions, the results indicate that this first prototype implementation of spatial collision avoidance already shows some desired results: during the mazes with spatial sound feedback the test subjects collided only half as often with walls than they did in the maze that provided collision notification only (the "thump" condition), shown in Figure 6b. This result supports hypothesis **H1** we made in section 4.

After their first collisions, subjects generally showed signs of approaching walls and corners more carefully, probing forwards with several small movements, which in the collision avoidance conditions provided them with more information about their wall distance than in the "thump" condition, where no feedback was given before the sudden impact. With collisions becoming fewer in the later conditions, collision fear sank moderately, as the questionnaire results in Figure 7f show. However, this could also in part be because of training after traversing several mazes; to clarify this a more comprehensive user study with more participants is necessary, where different groups of subjects are assigned different conditions, or complete the conditions in different orders. Also, in informal conversation after finishing the test some participants said that while their immediate fear of sudden unexpected collision sank after the spatial feedback started, they became more constantly aware of the presence of the walls (in itself a desired effect!) and the impending collision danger they posed, leading to a kind of constant background "fear" level that they did not feel in the (CN only) "thump" condition.

6.2 Wall distance

The differences in wall distance are only slight, but "chord" already yielded a significant difference to the initial "thump" condition; however, the difference was not more than 5cm (in a 1.5m wide corridor). Interestingly, the "chord" sound was rated the most pleasant, most calming (least stimulating), and least deterring sound (neutral rating between deterring and attracting) of the three collision avoidance sounds used. As this sound led to at least equal, partly slightly better results than the other two collision avoidance conditions, this may point toward more neutral sounds with a calming and reassuring effect on users (shown for "chord" through the low collision fear rating) being better suited for subliminal CA feedback. This speaks against hypothesis **H3**, which assumed that the more deterring and unpleasant a sound is, the better it keeps users away from the walls.

With the "chord" condition resulting in a small, but nonetheless statistically significant improvement on wall distance compared to the "thump" condition, we can cautiously affirm our second hypothesis (**H2**), in which we expected that spatial collision avoidance feedback aids in staying on a path away from the walls, at least for this condition. To support it more strongly, we suggest that further tests are necessary that offer a more spacious environment.

6.3 Helpfulness of collision avoidance

The questionnaire results regarding the helpfulness of collision avoidance feedback point to the same conclusion as the collision frequency and wall distance results — that the aural feedback is helpful. Although the differences between the collision avoidance conditions and the "thump" condition are not pronounced enough as to be statistically significant, they can be seen as an indication that we are going in the right direction with the design of the spatial feedback.

6.4 Completion time

The difference in completion time is not easy to interpret. Average completion time was 10 seconds faster in the second ("buzz") and fourth ("chord") conditions than in the first ("thump") and third ("bubble"). Our current interpretation is that familiarization with the environment lead to faster maze traversal in the second maze ("buzz") compared to the first ("thump"), but the very active, uneven sounding "bubble" sound irritated subjects and delayed them. The calming "chord" condition again resulted in a faster navigation.

7. CONCLUSIONS

The results from the initial user study indicate that the present prototype for the collision avoidance concept already achieves some of the goals set at the beginning of the project. Providing CA feedback through spatial sound leads to fewer collisions with walls and an increased awareness of the walls surrounding the user. It did not have a strong effect on adherence to the ideal, middle path in the test environment. This is partly due to the very confined and narrow nature of the maze environments we used, although there were small improvements in average wall distance detected. Here lies an opportunity to refine the feedback and study its effects in environments of varying narrowness or openness, to gain more insights into the benefits of spatial obstacle avoidance feedback.

Users reacted positively towards the spatial feedback, but did not always understand its purpose correctly. Also, they did not perceive an increased realism of their environment when provided CA feedback (although they did not perceive a *decreased* realism either). The successfully achieved effect of making them more aware of their surrounding geometry was seemingly countered by unsure reactions to the unfamiliar nature of the feedback. Especially persons that often played computer games felt that the sounds mostly did not fit the visuals, which seems to be an important factor for the acceptance of the soundscape and the world it helps to evoke.

While being conceived as a subtle and subconscious navigation aid, in its present form, our CA feedback was still very noticeable to the user and was thus perceived and interpreted more consciously; a way to counter this effect could be to integrate the CA feedback into a general ambient soundscape, from which it could gradually emerge when the danger of collision rises. Further user studies could compare the performance of user groups that have or have not been told beforehand what the purpose of the spatial sound is, and if this would change their perceptions of the soundscape surrounding them.

8. FUTURE WORK

Based on the current results, there are several possibilities for further user studies and improved collision avoidance system design:

The relative closeness of average wall distances in the narrow maze suggests that the observation of user paths through more spacious environments would possibly yield more pronounced differences in wall distance, and thus improve the evaluation of the effectiveness of different sound (or sound parameters).

User responses indicate that combining the sounds with fitting visuals would enhance the perception of reality, and fit the mental model better when trying to make sense of the sounds. Comparing several kinds of collision avoidance feedback sounds with fitting versus unfitting visuals of the surroundings, or wall textures, could yield more pronounced insight into the effectiveness of sound-enhanced collision avoidance techniques.

Another possibly interesting subject for further study would be a juxtaposition (or even combination) of wall-deterring collision avoidance feedback (as currently implemented) with some kind of inverse, optimal-path-marking positive feedback in the form of pleasant sounds.

Also, we envision the comparison of different kinds of feedback and sound modulation for indicating obstacle distance, including frequency modulation, equalizing, or cross-fading different versions of sounds, as well as using floor rumbling as a proximity indicator, adding an additional haptic modality to the current audiovisual feedback. We implemented prototypes for some forms of such additional feedback, like using music and amplitude-modulated floor rumbling for obstacle distance notification, but these would have to be refined and tested, which exceeded the scope of the presented initial study.

Finally, we already had test users who were familiar with the CA feedback successfully traverse the labyrinth blind (with projection turned off), although it took them a relatively long time to do so, especially probing for the direction of the path at corners. Comparative experiments could be made with groups of sighted, blind-folded, and blind subjects navigating the labyrinth with sound only.

9. ACKNOWLEDGMENTS

We would like to thank Matthias Haringer for supporting our work with the sound server, and Kristopher Blom for general technical guidance during the implementation phase of the CA feedback project. We also thank all test subjects taking part in the various user studies.

10. REFERENCES

[1] C. Afonso. Sound-enhanced spatial awareness for navigation in virtual environments. Diploma thesis, University of Hamburg, Department of Informatics, im.ve Group, 2011.

[2] D. Aguerrevere, M. Choudhury, and A. Barreto. Portable 3d sound/sonar navigation system for blind individuals. In *2nd LACCEI*, Miami, FL, 2004.

[3] K. J. Blom and S. Beckhaus. Virtual collision notification. In *Proceedings of the 2010 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 35–38, Waltham, MA, USA, 2010. TechNote.

[4] T. Garner, M. Grimshaw, and D. A. Nabi. A preliminary experiment to assess the fear value of preselected sound parameters in a survival horror game. In *Proceedings of the 5th Audio Mostly Conference: A Conference on Interaction with Sound*, AM '10, pages 10:1–10:9, New York, NY, USA, 2010.

[5] A. Gonot, S. Natkin, M. Emerit, and N. Chateau. An experiment in the perception of space through sound in virtual world and games. In *Proceedings of CGAMES*, 2006.

[6] N. Hess, J. D. Wischweh, K. Albrecht, K. J. Blom, and S. Beckhaus. ACTIF: An interactor centric interaction framework. In *Proceedings of the ACM conference on Virtual Reality Software and Technology (VRST) 2008*, pages 39–42. ACM Press, 2008.

[7] S. Holland, D. R. Morse, and H. Gedenryd. AudioGPS: Spatial audio navigation with a minimal attention interface. *Personal and Ubiquitous Computing*, 6(4):253–259, September 2002.

[8] T. Ifukube, T. Sasaki, and C. Peng. A blind mobility aid modeled after echolocation of bats. In *IEEE Transactions on Biomedical Engineering*, volume 38, pages 461–465, May 1991.

[9] P. Meijer. An experimental system for auditory image representations. In *IEEE Transactions on Biomedical Engineering*, volume 39, pages 112–121, 1992.

[10] S. Shoval, J. Borenstein, and Y. Koren. Mobile robot obstacle avoidance in a computerized travel aid for the blind. In *IEEE Robot. Autom. Conf. 1994*, pages 2023–2029, San Diego, CA, May 1994.

[11] M. Supa, M. Cotzin, and K. Dallenbach. "Facial vision": The perception of obstacles by the blind. *The American Journal of Psychology*, 57(2):133–183, 1944.

[12] T. Tullis and B. Albert. *Measuring the User Experience*. Morgan Kaufman, Burlington, MA, 2008.

[13] V. Twersky. On the physical basis of the perception of obstacles by the blind. *The American Journal of Psychology*, 64(3):409–416, 1951.

[14] B. N. Walker and J. Lindsay. Navigation performance with a virtual auditory display: Effects of beacon sound, capture radius, and practise. *Human Factors*, 48:265–278, 2006.