Collision Avoidance in Virtual Environments through Aural Spacial Awareness

Christian Afonso
im.ve
Department of Informatics
University of Hamburg
Vogt-Kölln-Str. 30
22527 Hamburg
2goncalv@informatik.uni-hamburg.de

Steffi Beckhaus
im.ve
Department of Informatics
University of Hamburg
Vogt-Kölln-Str. 30
22527 Hamburg
steffi.beckhaus@uni-hamburg.de

CHI 2011, May 7–12, 2011, Vancouver, BC, Canada.
ACM 978-1-4503-0268-5/11/05.

Abstract

In this paper we describe a new technique to make users aurally aware of walls surrounding them in a Virtual Environment (VE). This Collision Avoidance (CA) technique improves upon familiar Collision Notification (CN) feedback by constantly informing the user of his proximity to his surroundings through the playback of directional sounds. To render the aural CA feedback we use spatial sound played over surround loudspeakers, in addition to haptic feedback from a vibrating sound floor to signify collisions.

Keywords

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

ACM Classification Keywords

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

General Terms

Design, Experimentation, Human Factors, Performance

Introduction

Collision Notification is used in all kinds of Virtual Environments, ranging from games to scientific simulations. Upon colliding with the surroundings, a user receives feedback, such as a loud thump sound [2] or haptic force-feedback, depending on available equipment and the implementation. The experience of the unpleasant collision response might subsequently motivate the user to avoid further collisions.

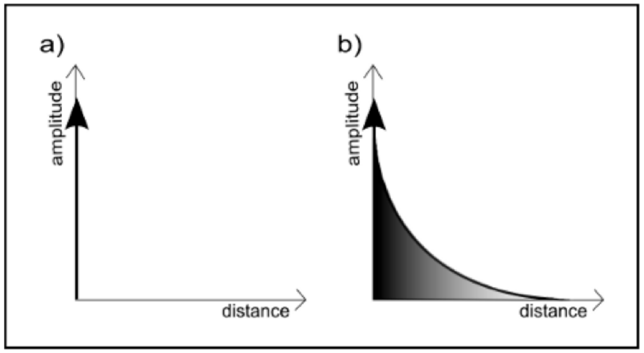


figure 1. Collision Notification (a) and Collision Avoidance (b). Depicted is the feedback given to users depending on the distance to a wall (positioned at distance 0). The amplitude denotes the strength of the feedback.

However, until the moment of collision the user will not be alerted to the imminent danger of impact (fig. 1a). In environments where 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 immersion is desired. We propose to shift from collision notification alone to tools for *collision avoidance*. The idea is, instead of letting users only know about their impact, to subconsciously guide them in such a way that they intuitively do not collide in the first place.

In real surroundings, we normally avoid objects intuitively. Virtual Environments (VE), however, often fail to give the user a full sense of presence in the environment. One of the reasons may be that, in reality, subtle non-visual cues exist that currently have no counterpart in VEs. Examples are temperature difference (cold exterior walls radiating coolness in an otherwise warm room) or changing acoustic characteristics (increase in early reflections near walls). Such cues mostly are not consciously perceived, yet may provide a subconscious navigation aid in reality.

Aural Collision Avoidance feedback in VEs would reinsert some degree of multimodal awareness of a user's surroundings, as she could both see and hear them. This might result in an increased navigational confidence due to the implicit subtle notification about a suitable distance, thus a certainty that no surprising sudden collisions will occur.

Therefore, we designed and implemented a collision avoidance system that uses spatial sound to notify the user of close obstacles (usually walls). Directional warning sounds increase in volume when the user gets closer to surrounding geometry (fig. 1b). We conducted a user study to compare user navigation through narrow maze-like maps with and without spatial proximity sound feedback.

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 (In [3] Gonot et al. compare *contextualized* and *decontextualized* beacons). Walker and Lindsay [7] studied how different kinds of sounds, different capture radius, and practise affected the adherence of test users to a path connecting localized beacon sounds. The *AudioGPS* implemented by Holland et al. [5] 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 [2] 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.

Apart from the motivation to extend existing Collision Notification techniques, we drew inspiration from real-world applications designed to help blind persons aurally perceive their surroundings. One of these systems is the *vOICe* [1]*,* 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. [6] 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.

System Design and Implementation

Our prototype system consists of modules for the ACTIF interaction framework [4] and can be integrated into existing VR applications. The ACTIF framework provides an abstraction for the steps of interaction processing, ensuring modularity and easy interchangeability of modules. Input data from the interaction devices (such as input buttons on a joystick/wand or positions of body tracking markers) is used to change the state of a *Virtual Actor* (VA), an abstract representation of the user in the framework without visual representation. The changes are interpreted by *Interpretation Modules* and applied to the scene graph by *Consequence*s. We use OpenSceneGraph[1] for graphics rendering and VR Juggler[2] for controlling input and output devices. As our display device we employ an "L-Shape", consisting of two projection screens, one in front of the user, one on the floor. For stereoscopic rendering we use "realD" shutter glasses synchronized with the stereoscopic projectors. Our sound setup consists of four speakers, arranged around the projection area (see the sketch in fig. 3), and low-frequency speakers built into the floor for haptic feedback. The user navigates through the VE using an omnidirectional wand device (Nintendo Wiimote fitted with optical markers for 6DOF tracking; from the Wiimote only the buttons are used for navigation, not the motion sensing functionality) and through motion tracking by the ARTtrack system[3].

---

[1] http://www.openscenegraph.org/projects/osg

[2] http://www.vrjuggler.org/

[3] http://www.ar-tracking.de/

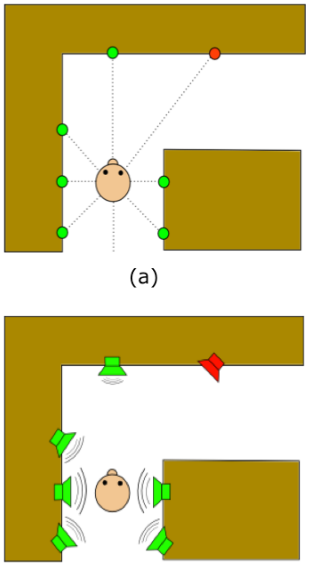


figure 2. Placement of the Sound Limbs (a) and Sound Sources (b).

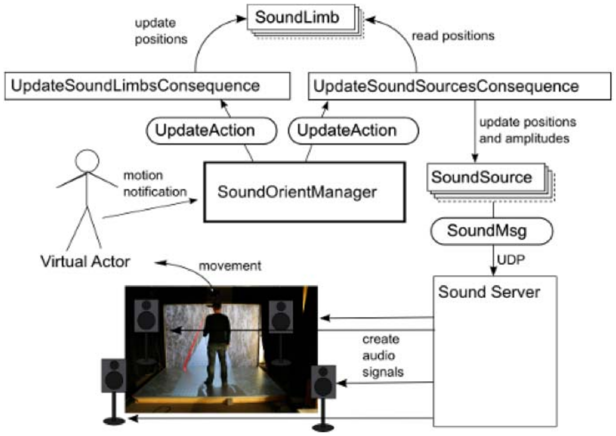


figure 3. Schematic diagram of the update loop of the collision avoidance system.

The distance sensors are realized as virtual "Sound Limbs" attached to the Virtual Actor (fig. 2a). Each time the user moves, the position of the VA is changed, and ray picking is performed in 8 equidistant directions around the Actor (The backwards-facing sensor is not seen in the image, as the wall in that direction is too far away). The Sound Limbs are positioned at the intersection points and store their distance to the Actor. If the distance is greater than a predefined maximum, the Sound Limb is turned inactive (indicated in red).

Subsequently, a virtual spatial sound source is placed at each Sound Limb position (fig. 2b). The amplitude of each sound is computed as a function of the stored distance of the associated Sound Limb; smaller distance yields increased amplitude (fig. 1b). For our user study, the function is chosen such that the sound is relatively low (but already audible), if the user stands in the middle of the narrow corridor, and reaches its maximum just before collision.

The centerpiece of the collision avoidance system is the manager module (see "SoundOrientManager" in fig. 3) that controls the interaction of the distance sensors and sound sources. The manager receives updates about changes to the Actor from the ACTIF framework's Core module. It sends Update Actions to the Sound Limb and Sound Source Consequences, which update the Sound Limb positions, the Sound Source positions and their amplitudes, respectively, as described in the previous paragraphs. The updated state of the Sound Sources is encapsulated into a *Sound Message* (one message for each source) and sent to the sound server, which in turn transforms them into physical audio signals for the speakers, resulting in 3-dimensional sound perception for the user standing in the virtual environment.

User Study

We conducted an initial user study to measure the effect of the Collision Avoidance system on the navigation performance of test users. We also compared it with the effect of Collision Notification alone. The VE employed for the tests consisted of various maze-like environments (although only consisting of one single possible path) with narrow corridors of 1.5m width (fig. 4), to make collision likely.

We tested four conditions, each corresponding to one maze: One pass with CN feedback only (a loud "thump" was heard and felt from the floor upon collision), and three passes with different sounds played as CA feedback (an electric buzz, a deep synthetic bubbling, and a continuous minor chord), combined with a loud electric buzz and a thump from the floor as CN.

The task presented to the users was to move through each maze swiftly from the starting point to a portal at

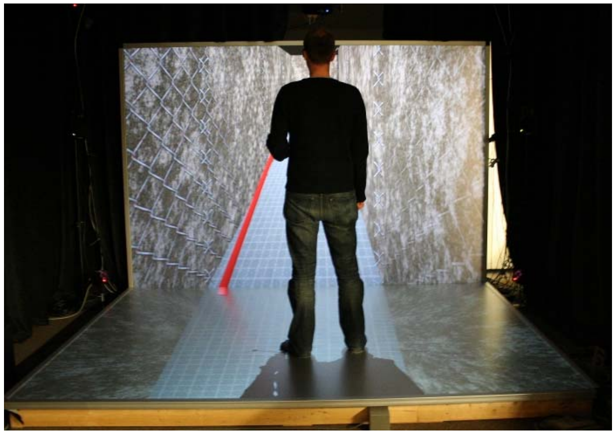


figure 4. User standing inside the maze. The red ray is attached to the wand device, indicating movement direction.

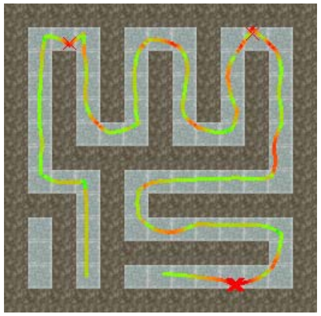


figure 5. Example path log through a maze. The color of the path signifies the user's distance to the walls, from light green (optimal distance) to red (very close to the wall). Red X's mark collisions.

the end. Users moved inside the VE by pressing buttons for translation and rotation on the Wiimote and pointing the Wiimote in the desired direction simultaneously. After each maze, the test users answered a short questionnaire. Their movements through the mazes were recorded in log files and corresponding path visualization images (see fig. 5 for an example).

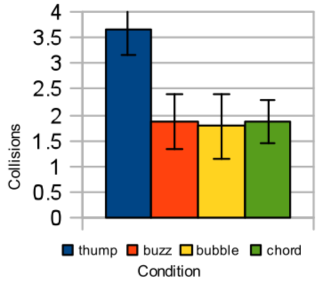
First Results

The initial study with 19 participants already shows promising results. We performed paired-sample t-tests on the log data to compare the results of the different conditions. When being provided with full Collision Avoidance feedback, on average a user collided only half as often with the walls than with Collision Notification feedback only (fig. 6a). In addition, users indicated in the questionnaires that all three types of CA feedback aided their navigation slightly more than the CN feedback only.

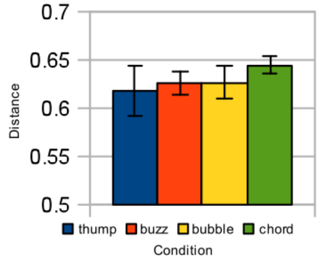
The average distances of users to walls in the different conditions, however, are relatively similar (fig. 6b). The mean wall distance was computed per user and condition, and an average was taken over the whole sample (as the corridors were 1.5m wide, the maximal possible wall distance was 0.75m). The only CA condition significantly different (with >99% confidence) from the “thump” condition is the one with the musical chord, and, even here, the difference is less than 5 cm.

Looking at the average completion time (fig. 6c), the differences between certain conditions are a bit more pronounced: Statistically, “thump” and “buzz”, “bubble” and “buzz”, and “bubble” and “chord” are significantly different (with 99% confidence). Our current interpretation is that generally, CA feedback speeds the navigation through the maze if the sound is simple and even (and not too active, like “bubble”). However, to interpret the meaning of these differences with some certainty, more detailed tests will have to be done.

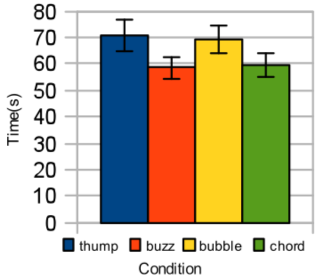
User responses from the questionnaires indicate that the fear of imminent collision was slightly reduced in the latter CA conditions, though informal conversations after completion of the test indicate that reduced fear resulting from the knowledge that collision was not imminent was offset by increased collision fear resulting from the constant aural awareness of the walls (and the possibility of collision with them). The buzz sound was considered the most unnerving and annoying. The musical chord, yielding the lowest collision fear level, was mostly perceived as relaxing and comforting, though some users found it rather discomforting due to its monotony.



(a) Average number of wall collisions.



(b) Average of the per-user mean wall distance (in m).



(c) Average completion time (in s).

figure 6. Evaluation results. Conditions are (from left to right): thump (CN only), buzz, bubble, and chord (CA+CN).

Users commented that a coherent coupling of fitting sounds and images greatly enhanced their perception of their surroundings as realistic (in this case, coupling the electric buzz sound with an alpha-textured see-through wire-fence texture applied to the walls) and aided the correct perception of the spatial sound as coming from the walls; without visuals fitting the sounds (like hearing a buzz but only seeing stone walls) they sometimes perceived the sound as coming from somewhere else in the maze (behind the walls), tried to find the source, and were confused about the sound changes not corresponding to their mental model.

Future Work

Based upon our current results, there are several possibilities for further studies and improved CA system design. The relative closeness of average wall distances in the narrow maze suggests that the evaluation of user paths through more spacious environments could yield more pronounced differences in wall distance.

User responses indicating stronger perception of reality when combining sounds with fitting visuals poses the possible benefit of comparing several kinds of CA feedback sounds with fitting/unfitting surroundings or wall textures. Different kinds of sound modulation for indicating obstacle distance can be compared, 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.

Another possibly interesting subject of further study would be a juxtaposition (or even combination) of the described wall-deterring CA feedback with some kind of inverse, optimal-path-marking positive feedback in the form of pleasant sounds. We currently work on an open space VE that aids navigation and, at the same time, subtly guides users through this environment. This is a form of *guided exploration* that can be utilized in sound only environments as only modality or in full VEs as subtle supplement to other modalities.

The principles of the technique can be applied to real-world applications like navigation aids for the blind. For this, the omnidirectional ray-picking in the scene graph is substituted with physical distance sensors (e.g. sonar/ultrasound, or optical depth recognition devices), and the spatial sound rendered on head-phones using head-related transfer functions (HRTFs).

Citations

[1] Augmented Reality for the totally blind: The vOICe http://www.seeingwithsound.com/

[2] Blom, K.J., Beckhaus, S. Virtual Collision Notification. *Proc. 3DUI* (2010), TechNote, 35–38.

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

[4] Hess, N., Wischweh, J.D.S., Albrecht, K., Blom, K.J., Beckhaus, S. ACTIF: An Interactor Centric Interaction Framework. *Proc. VRST* (2008), 39-42.

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

[6] Ifukube, T., Sasaki, T., Peng, C. A blind mobility aid modeled after echolocation of bats. *IEEE transactions on Biomedical Engineering 38(5)* (1991), 461-465.

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