Multi-User Interaction in Virtual Audio Spaces

Florian Heller

RWTH Aachen University

52056 Aachen, Germany

flo@cs.rwth-aachen.de

Thomas Knott

RWTH Aachen University

52056 Aachen, Germany

thomas.knott@rwth-aachen.de

Malte Weiss

RWTH Aachen University

52056 Aachen, Germany

weiss@cs.rwth-aachen.de

Jan Borchers

RWTH Aachen University

52056 Aachen, Germany

borchers@cs.rwth-aachen.de

Abstract

Audio guides are a common way to provide museum visitors with an opportunity for personalized, self-paced information retrieval. However, this personalization conflicts with some of the reasons many people go to museums, i.e., to socialize, to be with friends, and to discuss the exhibit as they experience it [1]. We developed an interactive museum experience based on audio augmented reality that lets the visitor interact with a virtual spatial audio soundscape. In this paper, we present some new interaction metaphors we use in the design of this audio space, as well as some techniques to generate a group experience within audio spaces.

Keywords

augmented reality, museum, spatial audio, mobile computing

ACM Classification Keywords

H.5.1. [Information interfaces and presentation] Multimedia Information Systems --- Artificial, augmented, and virtual reality; H.5.3 [Information interfaces and presentation] Group and Organization Interfaces – Collaborative computing, synchronous interaction system.

Introduction

The goal of CORONA is to create an interactive audio experience in the Coronation Hall located in the city hall of Aachen, Germany. There are only few visual cues in this 45m x 20m large site that indicate what important ceremonies took place in this hall and, as the hall is still used for festivities, it is not possible to install exhibits that would overcome this problem. Also, the visual impression of this historical site was not to be modified, so we decided to augment the physical space with a virtual audio space.
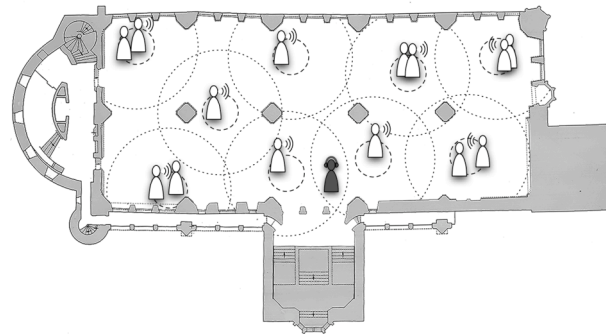

Figure 1: floor plan of the Coronation Hall with the audio space and its source areas (small circles).

This audio space is an acoustic simulation of a virtual scene, rendered on a mobile device handed out to the visitors. Our mobile device consists of an Apple iPhone extended with some additional hardware (cf. Figure 3) for the localization process. The availability of the spatial audio rendering framework OpenAL [7] on the iPhone makes it an optimal platform for such an experience. To determine the user's position in the hall, we use the Ubisense real time location system [8]. As

we also need the head orientation, we mounted a compass sensor on the headphone. The rendering process takes into account the position of the virtual sound sources and the position and head orientation of the visitor to calculate an imitation of natural sound propagation. The visitor wears headphones and experiences an impression, e.g., of someone standing a few meters away and singing. The visitor explores this virtual audio space by moving in the physical world. The more distant a sound source is, the more its audio stream fades from clear speech to murmur. This technique allows the visitor to explore the virtual audio space without any representations in the physical world, just by listening and moving in the direction of an interesting sound. We call this navigation by ear. By manipulating the acoustic model of our virtual space, we are also able to, e.g., fade out the signals of distant sound sources thus avoiding a disturbing mix of several signals that would lead to confusion.

Some preliminary tests with an optical tracking system revealed, that even with a very high tracking accuracy, the achievement of a satisfying result needs very elaborate tuning of the available audio rendering parameters. The requirements on the location tracking systems are very tough, but we also have to consider the special environment the system is installed in. Our decision is therefore a tradeoff between accuracy and visual unobtrusiveness.

The audio space consists of about ten source areas and a background atmosphere. A source area, which is the basic entity for the visitor, can either be a single character that just tells the visitor a story or several characters talking among themselves. These areas consist of two nested zones. The outer zone describes

where the sound of the source is audible, whereas the inner one describes the area where the voice of the virtual character is clearly understandable (cf. Figure 2). The transition between clear speech and background atmosphere is continuous to ensure that, as suggested in [9], the transition between immersion and non-immersion is fluid and thereby does not confuse the visitors. However, even in the calm areas the background atmosphere should still be audible to reassure the visitor that the system is still working [11].

The Audio Space

As scenery for our audio space, we decided to use a medieval coronation feast, a historical event that took place many times in this hall. It includes a set of characters that can convey different types of information. A similar concept is analyzed in [6], with the result that historic characters are well suited to vividly communicate historic data. History becoming alive is also reported as an important factor for immersion and enjoyment in [9]. By illustrating facts from different points of view, information is communicated in an easily understandable way, e.g., a discussion between servants from the 15th century will reveal conflicts and insights related to their social class.

Visitors will not be part of these conversations but will instead be observers of the ceremony; they are able to listen to the conversation of the virtual characters without interruptions by the guide. Even if audio augmented reality has already been deployed on a small scale in other museums [5], most of our visitors will have no experience with virtual spaces. Thus, it is crucial to keep the interaction simple.

However, a virtual character of the ceremony can engage the visitor in some activities, e.g., follow him to a special place. To not draw the attention away from the audio scene and the Coronation Hall, the mobile device does not serve as controller for the audio space. The visitor can thus leave the guide in her pocket [9].

From preliminary interviews we know that there are two main tendencies in visitor behavior. The first one decides from an overview which exhibits are interesting. The other approach is to follow the designated tour through the museum. Our design supports both of these behaviors. To get an overview, the visitor can walk through the entire space and shortly listen to the different conversations. Finding the interesting points again is easy because they stay at a fixed position and thereby can be memorized spatially.
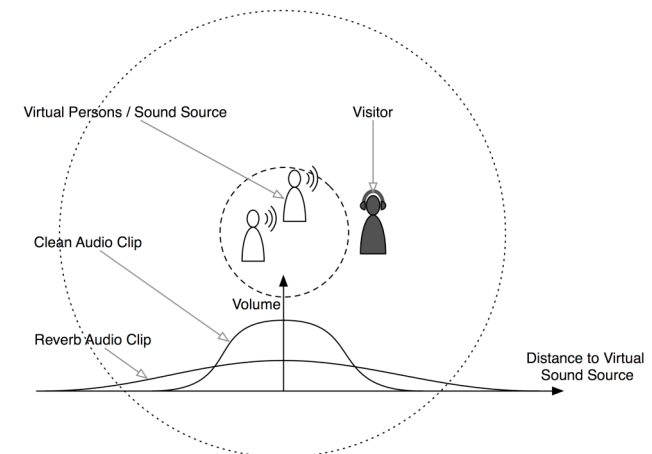

Figure 2: The design of a source area. In the outer circle the voices of the two talking characters become clearer towards the center.
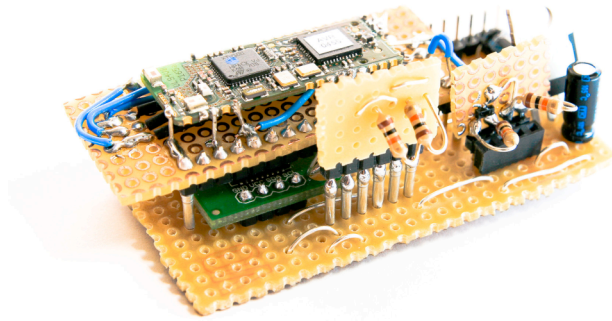
Figure 3: A prototype of the headphone module with tracking hardware and compass sensor.

Group Experience

For the single visitor, the interaction has been designed to be as natural as possible using the virtual character metaphor. Nevertheless, by requiring the visitor to wear headphones, we isolate her from the group she might be with in the Coronation Hall. As already analyzed in [1,2], people often come to museums for social factors, i.e., being with friends and discussing an exhibit as they experience it. To minimize the isolation on the hardware side, we use lightweight open headphones that dampen the outside signals only a bit, such that talking to others is still possible while wearing the headphones. However, we also have to provide features that attenuate the isolation and allow groups to experience the audio space together.

Shared Audio Space

A group of two or more visitors has some additional requirements over the solo visitor. For example, before starting a conversation, it is essential to know if the desired partner is currently listening to an explanation or not. It is also important to know if he already heard a specific part of the explanation. To take into account the special interaction within such a group of visitors, we decided to let them share a common audio space instead of experiencing a personalized one. Therefore, playback of audio samples is synchronized on all mobile devices.

The experience of the audio space as a group is realized by physical proximity. If a visitor wants to listen to the scene with someone else, she has to stay close. However, the sample is still rendered according to the individual position of each visitor.

Our design supports the activity patterns described in [2], although we do not provide an eavesdropping functionality. For *shared listening* and *following*, the visitors just have to walk through the exhibition together. To *check-in* to another visitor, just walk to the corresponding person. However, to know what another visitor has already heard, one just has to observe the path a certain visitor follows in the hall.

Inter-Visitor Interaction

Users of a similar system in a Zoo reported that the sharing of a common audio space had a stimulating influence on their communication with each other [3]. To facilitate a conversation, some places in the audio space remain calm such that we can rely on the ability of the human hearing to selectively focus on specific conversations in a mixed environment, known as the cocktail party effect [10].

Observations

In a preliminary Wizard-of-Oz study, seven users interacted with a small version of the CORONA audio space. The scenario contained two main source areas, a speaker and a band, and some atmospheric background. The users were asked to explore the space and to designate where the source areas are located. We observed that turning the head is the key to navigation by ear. The localization task was accomplished very well but users reported that it is more difficult to localize the band than the speaker. In general, even though it was not a realistic experience, our users liked the playful approach.

Outlook

We developed a flexible hardware platform for immersive audio experiences. Besides the effort to install and calibrate the location tracking system, it can be adapted to a lot of different usage scenarios. The integration of the compass as key information source, and the audio rendering process into the mobile device reduces latency and makes the system scalable. Although augmented audio reality systems have already been deployed on a small scale in other museums [5], most of our users are first time users and probably never got in touch with augmented reality. Therefore, we tried to keep the interaction metaphor as simple as possible. However, the number of possible audio augmented scenarios is large. A 3D audio first person shooter game has successfully been tested [11].

Our preliminary studies show that users are willing to hand-off the playback control to the system. However, in some situations playback control might be useful, e.g., a father wants to play a specific passage from the discussion again to help his children understand the information. But, this decreases the immersive effect and makes the interface more complex. The possibilities of playback control have to be analyzed carefully.

The simple audio rendering mechanism we use does not create a realistic audio experience. However, it is plausible enough to allow directed navigation by ear. A point we want to address with this project is, how the realism of the spatial audio rendering process affects the interaction with the virtual audio space.

Acknowledgements

References

[1]   Bederson, B.B. Audio Augmented Reality: A Prototype Automated Tour Guide. *Conference companion CHI 1995*, ACM Press (1995).

[2]   Grinter, R.E., Aoki, P.M., Hurst, A., Szymanski, M.H., Thornton, J.D., and Woodruff, A. Revisiting the Visit: Understanding How Technology Can Shape the Museum Visit. *Proc. CSCW 2002*, ACM Press (2002), 146-155.

[3]   Stahl, C. The Roaring Navigator: A Group Guide for the Zoo with Shared Auditory Landmark Display. *Proc. MobileHCI 2007*, ACM Press (2007), 383-386.

[4]   Woodruff, A., Aoki, P.M., Hurst, A. and Szymanski, M.H. The Guidebook, the Friend, and the Room: Visitor Experience in a Historic House. *Ext. Abstracts CHI 2001*, ACM Press (2001), 273-274.

[5]   Terrenghi, L. and Zimmernmann, A. Tailored Audio Augmented Environments for Museums. *Proc. IUI 2004*, ACM Press (2004), 334-336.

[6]   Dow, S., Lee, J., Oezbek, C., MacIntyre, B., Bolter, J.D. and Gandy, M. Exploring spatial narratives and mixed reality experiences in Oakland Cemetery *Proc. ACM SIGCHI International Conference on Advances in computer entertainment technology 2005,* ACM Press (2005), 51-60.

[7]   OpenAL Project Website http://www.openal.org, 2009

[8]   Ubisense RTLS http://www.ubisense.net, 2009

[9]   Reid, J., Geelhoed, E., Hull, R., Carter, K. and Clayton, B. Parallel Worlds: Immersion in Location-Based Experiences *Ext. Abstracts CHI 2005,* ACM Press (2005), 1733-1736.

[10] Arons, B., A Review of the Cocktail Party Effect. *Journal of the American Voice I/O Society*, 12-7, 1992, pp. 35-50*.*

[11] Reid, J., Cater, K., Fleuriot, C. and Hull, R. Experience Design Guidelines for Creating Situated Mediascapes, *Mobile and Media Systems Laboratory,* HP Labs, 2005

[12] DEMOR, A Location Based 3D Audio Game http://student-kmt.hku.nl/~g7/site/index_.html, 2004