Designing Audio-Augmented Environments

Andreas Zimmermann and Lucia Terrenghi

Fraunhofer Institute for Applied Information Technology, Schloss Birlinghoven,
53754 Sankt Augustin, Germany

Abstract. The paper deals with the design of an intelligent user interface augmenting the user experience in a museum domain, by providing an immersive audio environment. We highlight the potential of augmenting the visual real environment in a personalized way, thanks to context modeling techniques. The LISTEN project, a system for an immersive audio augmented environment applied in the art exhibition domain, provides an example of modeling and personalization methods affecting the audio interface in terms of content and organization. In addition, the outcomes of the preliminary tests are here reported.

1 Introduction

One core idea of ubiquitous computing [13] is that items of our daily environments will acquire computational capacities, e.g. sensing data in the context, elaborating them and adapting the system to the context of interaction, will provide new functionalities and enable new activities. The physical reality will be overlaid by an additional virtual layer (Mixed Reality) and by naturally moving in the space and/or by manipulating physical objects in our surroundings we will act upon information in the virtual layer. This enables new augmented user experience and funds the basis for new challenges in user interface design.

The LISTEN project [7] conducted by the Fraunhofer Institute for Media Communication is an attempt to make use of inherent “everyday” integration of the aural and visual perception [3]. In October 2003 this system was applied for the visitors of the August Macke art exhibition at the Kunstmuseum in Bonn, in the context of the “Macke Labor” [12]. The users of the LISTEN system move in the physical space wearing wireless headphones, which are able to render 3-dimensional sound, and listen to audio sequences emitted by virtual sound sources placed in the environment. The visitors of the museum experience personalized audio information about exhibits through their headphones. The audio presentation is adapted to the users’ contexts (i.e. interests, preferences, and motion), providing an intelligent audio-based environment [10].

Museum and exhibitions as domains have already been explored in several research projects. Many of them focus on the goal to provide a museum guide including content concerning the artworks [8], to immerse the user in a virtual augmented environment built in virtual museums [1], or to provide orientation and routing functionalities

Zimmermann A. and Terrenghi L. (2004).
Designing Audio-Augmented Environments.
In *Proceedings of the 1st International Workshop on Ubiquitous Computing*, pages 113-118
DOI: 10.5220/0002673201130118

[2]. The LISTEN project takes the challenge to provide a personalized immersive augmented environment, which goes beyond the guiding purpose, and is based on the combination of aural and visual perception. The viewers become inter-actors as soon as they move in real space. Thus, the physical behavior of the user determines the interaction. This dissertation focuses on the design and adaptation of this user interface, and serves as a proof of concept for our approach.

2 The Combination of Content and Context

While using the LISTEN system, users automatically navigate an acoustic information space designed as a complement or extension of the real space. The selection, presentation and adaptation of the content of this information space take into account the user's current context. For the flexible and easy combination of the appropriate content and the current context, we chose to provide a centralized Java based context and user modeling server. In addition, we use XML as the modeling language, which allows for a intuitive design, definition and configuration of the system components. Since this design tool will be used in various follow-up projects, we introduced a new model for the design of context-aware systems (cf. Figure 1).

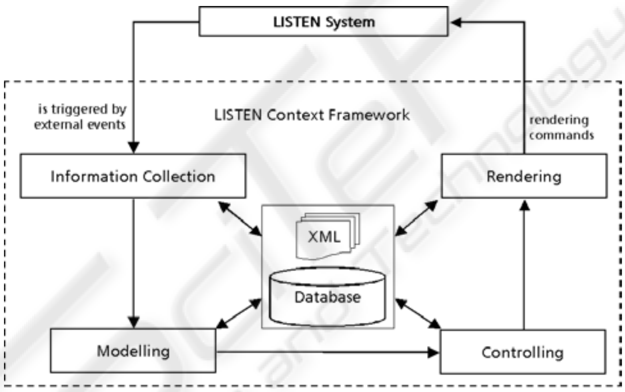

Fig. 1. The Context-Aware Architecture of the LISTEN System

In the following we describe the four steps of this abstracted model: information collection, modeling, controlling and rendering. In addition, we show the instantiation of this model by the implementation of the LISTEN at the museum (cf. Figure 2).

2.1 Information Collection

A sensor network placed in the environment monitors variable parameters and recognizes changes within the environment. A sensor server receives all incoming events (i.e. user interactions) sent by the application. These event descriptions are pushed into a database. Thus, an event history for every user is saved and an implicit user profile is recorded. The LISTEN system allows monitoring of three sensor values only: the user's spatial position and head orientation [4] and the time.

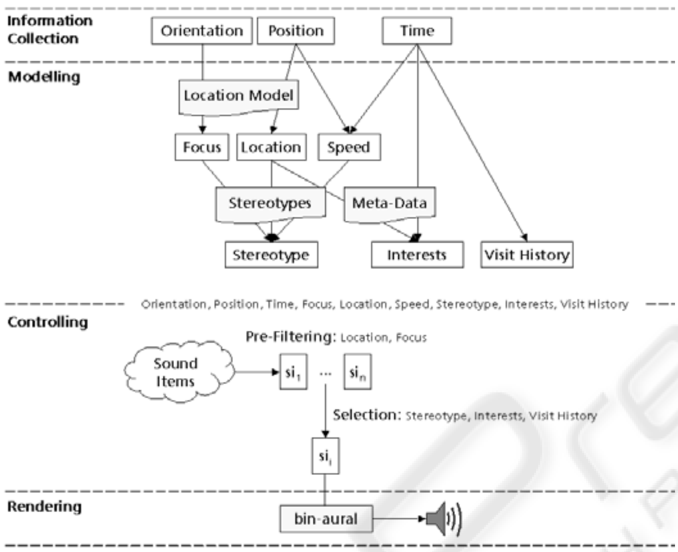

Fig. 2. The Instantiation of the Abstract Model

2.2 Modeling

In a second step the incoming sensor values are interpreted by several algorithms (e.g. machine learning or data mining algorithms) based on different models (e.g. overlay, statistic, etc.). Thus, semantically enriched information is extracted and more significant knowledge is gained relating to the user's behavior.

A location model, which segments the exhibit hall into *object zones* and *near fields* [5], allows the LISTEN system to interpret the user's position (i.e. location) and head orientation (i.e. focus). The most important variable within the user modeling process it the time. The analysis of the user's spatial position and the time results in the speed the user moves with. In addition, a combination of meta-information concerning the paintings and the time allows an assumption about the user's interests (e.g. the more time the visitor spends with a specific exhibit, the more s/he likes it) [9].

Furthermore, we chose to employ *stereotypes* to define the user's *observation type*. In contrary to [8] the user's classification into one stereotype is done manually by the user and cannot be changed automatically during runtime. Additionally, the personalization engine of the LISTEN application is not able to perform an automated clustering and derivation of new stereotypes.

The complete model of the context we use in LISTEN is represented by a set of attribute-value pairs.

2.3 Controlling

In a third step a controlling component is necessary to decide what consequences must be taken if certain conditions in the user's context and in the individual user model configuration appear together. Based on these information sources, the control component assembles a sequence of commands in order to adjust certain variable properties of the environment. Different sequences of commands lead to different kinds of information presentation.

A XML-configurable rule-system controls the sound presentation of the LISTEN scenery. First, a pre-filtering of sound-items is performed based on the user's current location and focus. From this list the best suited sound-item is selected referring to the user's visit history, stereotype and interests.

2.4 Rendering

Rendering means handling the connection back to the domain. This engine translates the assembled sequence of domain-independent commands into domain-dependent commands changing variable parameters of the domain. Thus, the decisions taken by the controlling component are mapped to real world actions.

Concerning the LISTEN system rendering means changing the audio augmentation. Basically this denotes a certain sound-item being played from a specific sound source. But moreover, this sophisticated auditory rendering process takes into account the current position and orientation of the user's head, in order to seamlessly integrate the virtual scene with the real one [6].

3 Adaptation Means

The basis for every kind of adaptation in LISTEN is the presentation of the sound space. Besides the selection of which sound item is to be played, other dimensions influence the sound presentation: for instance when, with which character (e.g. volume), from which direction, with which motion and how long a sound is played. With combinations of these possibilities, a wide range of adaptability is accomplished. Additionally, other means of adaptation are realized or planned.

3.1 Adaptation of the Space Model

In LISTEN users enter and leave zones in virtual space (cf. Section 2.2 and [11]). Some users want to step back and look at the object from a different viewpoint. Because the user still shows interest in this specific object, the associated zone should adapt to the user and expand up to a predefined point (zone breathing). Then the user is able to listen to the sound further on.

3.2 Adaptation to Social Context

By interpreting the time and the location of several users, the LISTEN system is able to identify clusters of people (spatially and temporally similar users). These people might want to receive the same audio information (e.g. like a family). Vice versa, breaking up clusters of people is also possible. This would lead to a better distribution of people among several objects.

3.3 Adaptation to the Level of Immergence

In LISTEN, interest in objects is expressed by the time a user's focus lingers on these objects. The complexity, the amount, and the style of already received information about one object corresponds to the level of interest. The sound presentation directly steps in the adequate level of information depth and style, if the currently seen exhibit complies with the user's interests.

3.4 Adaptation to Movement and Reception Styles

People walking through the environment show different kinds of common behavior (e.g. clockwise in museums [8]). Special attractor sounds emitted from different sound sources are used to draw the user's attention on certain objects. Thus, entire predefined tours through the environment are recommended (by artist, curator, personal interest) and dynamically adapted to the stereotypical type of movement and perception style.

4 Conclusions and Evaluation

In this paper we attempt to highlight the potential of tailored audio augmented environments design, a research field that needs to be further explored and still presents critical issues. The project here described provides a good example of the complexity of the modeling and design process of a tailored audio augmented environment.

The outcomes of first test at the Kunstmuseum in Bonn enabled a preliminary evaluation: impressions and refinement issues were brought out by 15 visitors, art curators, sound designers, and artists.

As a first positive result all the visitors enjoyed the combination of audio-visual perception and felt an augmentation of the environment. At the same time, both curators appreciated the possibility to deliver content concerning the artworks in an innovative, enriched, and less descriptive way.

As critical points eight testers mentioned problems to "feel" the interaction with the environment and four even reported that they were forced to approach the artwork very closely. These statements show the importance of a more dynamic space model and the necessity of landmarks in the virtual environment navigation, we are currently working on.

In addition, five users could not realize whether the changes in the audio virtual environment were due to their movements in the space or were part of the audio sequence. This misinterpretation of the user's spatial position in combination with the focus is currently under investigation.

Four users missed an explicit control channel for switching between different types of sound presentation. Since the LISTEN systems aims at a complete immersion of the users only through wireless headphones this issue will not be further pursued.

A main criticism in the evaluation was the use of stereotypes for content adaptation. The people do not like to be clustered and classified by their personal information consumption habits (fact oriented, emotional or overview oriented). The refined system will provide stereotypes that represent the user's moving styles like *sauntering*, *goal-driven* or *standing still*, because these stereotypes have less social impact, and are easy to detect and to revise.

References

1. Chitarro, L./Ranona, R./Ieronutti, L.: "Guiding Visitors of Web3-dimensional Worlds through Automatically Generated Tours". In: Proceedings of Web3D 2003: 8th International Conference on 3D Web Technology (2003).
2. Ciavarella, C./Paterno', F.: "Supporting Access to Museum Information for Mobile Visitors". In: 10th International Conference on Human-Computer Interaction. Crete, Greece (2003).
3. Eckel, G.: "LISTEN – Augmenting Everyday Environments with Interactive Soundscapes". In: Proceedings of the I3 Spring Days Workshop "Moving between the physical and the digital: exploring and developing new forms of mixed reality user experience", Porto, Portugal (2001).
4. Goiser, AMJ.: Handbuch der Spread-Spectrum Technik. Springer-Verlag Wien New York (1998).
5. Goßmann, J., Specht, M.: "Location Models for Augmented Environments". In: the Workshop Proceedings of Location Modelling for Ubiquitous Computing, Ubicomp (2001), 94-99.
6. Jot, Jean-Marc : Real-time spatial processing of sounds for music, multimedia and interactive human-computer interfaces. Multimedia Systems 7. Springer (1999).
7. LISTEN: http://listen.imk.fraunhofer.de. The LISTEN Website (2000)
8. Oppermann, R., Specht, M.: "A Context-sensitive Nomadic Information System as an Exhibition Guide". In: Proceedings of the Handheld and Ubiquitous Computing Second International Symposium, HUC 2000, Bristol, UK (2000), 127 – 142
9. Pazzani M. J./Billsus D.: "Learning and Revising User Profiles: The Identification of Interesting Web Sites". In: Machine Learning 27 (1997), 313-331.
10. Sawhney, N./Murphy, A.: "Designing Audio Environments – Not Audio Interfaces". Paper submitted to ASSETS '96, The 2nd ACM/SIGCAPH Conference on Assistive Technologies (1996).
11. Tramberend, H.: "Avango: A Distributed Virtual Reality Framework". In: Proceedings of the IEEE Virtual Reality '99 Conference, Houston,Texas, USA (1999).
12. Unnützer, P.: "LISTEN im Kunstmuseum Bonn, KUNSTFORUM International, Vol. 155 (2001), pp. 469/70.
13. Weiser, M.: "The Computer for the Twenty-First Century," Scientific American (1991), pp. 94-10.