LISTEN: Contextualized Presentation for Audio-Augmented Environments

Andreas Zimmermann and Andreas Lorenz
Fraunhofer Institute for Applied Information Technology
Schloß Birlinghoven
53754 Sankt Augustin, Germany
{Andreas.Zimmermann, Andreas.Lorenz}@fit.fraunhofer.de

Abstract

The paper deals with the awareness of and the adaptation to the context in audio-augmented environments. Taking into account the relationship between aural and visual perceptions, we focus on the issues and the potential of adapting intelligent audio interfaces to augment the visual real environment. The description of the LISTEN project, a system for the creation of immersive audio-augmented environments, is taken as showcase; in particular, our focus is on the modelling and personalization methods affecting the audio design and presentation. The outcomes of the preliminary tests are reported in order to describe the issues and complexity of intelligent user modelling in audio-augmented environments.

1 Introduction

The LISTEN project conducted by the Fraunhofer Institute in St. Augustin deals with the audio augmentation of real and virtual environments. The users of this system move in space wearing headphones and listen to audio sequences emitted by virtual sound sources placed in the environment. The audio pieces vary according to the user's spatial position and orientation of his/her head.

At the beginning of July 2003, a first LISTEN prototype was applied to an art exhibition at the Kunstmuseum in Bonn [20]. The visitors of the museum experience personalized audio information about exhibits through their headphones. In such a scenario, the distribution and authoring of this valuable information on exhibits is a non-trivial task [22]. The audio presentation takes into account the visitor's profile. Besides presentation, the system provides recommendations to the visitor regarding to his/her context. Recommended exhibition objects attract the visitor's attention by emitting sounds from its position.

Furthermore, the LISTEN environment offers the ability to adapt the audio presentation (i.e. the order of audio pieces, their content, and sound source) to the users' contexts (e.g. interests, motion, focus, etc.) and thus perform personalization. Since the user movement in physical space is the only interface, these movements have to be interpreted in order to derive a meaningful interest model of the users.

This paper aims at describing the main issues in the user modelling components that have been employed for the personalization of the LISTEN system. For this reason we have investigated three different approaches that seem to be suitable for our purposes: *Content-Based Filtering Systems* [e.g. [1,2]), *Collaborative Filtering Systems* (e.g. [6,15]) and *Context-Aware Systems* (e.g. [11,16,17]). Furthermore we examined hybrid systems (like [3,15]), which combine several approaches. They are often able to overcome some limitations of pure approaches and improve the quality of recommendations.

Section 2 of this paper gives an brief overview on context-aware systems discussing how context can be defined and giving an idea about what context is in the LISTEN system. A detailed description of the LISTEN system is provided in Section 3. The interpretation of the users' context and monitoring of the context's evolution over time answers the question, *why* should be adapted. Section 4 gives an answer to that question from a LISTEN perspective and focuses on *what* can be adapted within an audio-augmented environment. Thus, we describe possibilities of combining the order, the source (position and motion) and the content of audio pieces for adaptation purposes. Furthermore, we provide an overview on the first evaluation process of the LISTEN system in Section 5.

2 Context-Aware Systems

By definition, *context-aware systems* are aware of and adapt to the context of the user. Several approaches have defined context models and described different aspects of a context taken into account for context-aware systems. For example, Schilit et al. have mentioned [17]: where you are, who you are, and what resources are nearby. Dey and Abowd [9] discuss several approaches for taking the computing environment, the user environment, and the physical environment into account. Furthermore, they distinguish between primary and secondary context types: Primary context types describe the situation of an entity and are used as indices for retrieving second level types of contextual information. In this work we base our context modelling approach on the four dimensions of a context which Gross and Specht have considered in [11]:

Identity: The identity of a person gives access to the second level of contextual information. In some context-aware applications highly sophisticated user models hold detailed activity logs of physical space movements and electronic artefact manipulations and infer information about the user's needs, interests, preferences, knowledge, etc.

Location: We consider location as a parameter that can be specified in electronic and physical space. An artefact can have a physical position or an electronic location described by URIs or URLs. Location-based services as one type of context aware applications can be based on a mapping between the physical presence of an artefact and the presentation of the corresponding electronic artefact [10].

Time: Time is an important dimension for describing a context. Beside the specification of time in CET format categorical scales as an overlay for the time dimension are mostly used in context-aware applications (e.g., working hours vs. weekend). For nomadic information systems a process-oriented approach can be time dependent (similar to a workflow).

Environment or Activity: The environment describes the artefacts and the physical location of the current situation. In several projects, approaches for modelling the artefacts and building taxonomies or ontology about their interrelations are used for selecting and presenting information to a user. A group of people sharing a context is also part of the environment (social context), as well as the technical context describing which devices are used.

Context awareness enhances the possibility to design intelligent user interfaces: Their context dependency builds a bridge between user and system in order to improve the interaction usability. Smart modelling techniques are essential to acquire, represent and exploit context awareness. In particular, context modelling includes user modelling as a key issue. User models can be deduced explicitly or implicitly [12]. In the first case, the user states information explicitly, such as his/her interests, preferences, skills etc. In the latter case, inference techniques based on domain modelling, understanding the typical behaviour of similar users, and user interaction observation (tracking systems) are used in order to draw hypothesis about user interests, needs, and plans. Inference systems pose the challenge of defining what parameters should be measured and how, so as to acquire a complete and useful context description.

The interpretation of the overall context model may have impact on a large variety of adaptation targets like information presentation, content, modality definition, etc. Since the adaptation process is strictly dependent on the domain, the LISTEN specific process is described in Section 4.

3 Aural-Visual Interaction: The LISTEN Project

The goal of a 'Computer Augmented Environment' is to bring computational power to everyday contexts so as to enhance a seamless, invisible user interface. In order to augment and enhance the everyday life, rather than attempting to replace it, the computer systems need to be aware of the context [9]. The project we describe in this paper is an attempt to make use of inherent "everyday" integration of the aural and visual perception, developing an immersive audio-augmented environment.

The key idea of the LISTEN concept [8] is to place the individual perception of space at the centre of the interface design so as to convey an immersive user experience. By moving through real space users additionally navigate an acoustic information space designed as a complement or extension of the real space. Virtual acoustic landmarks play an equally important role than the visual ones for the orientation of the users in this augmented environment. Acoustic labels are attached to visual objects, thus affecting the soundscape and its perception. Fine-grain motion tracking is essential for the LISTEN approach because full auditory immersion can only be reached when the binaural rendering process takes into account the rotation of the user's head. 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.

A first prototype of the LISTEN application has been applied for the August Macke art exhibition at the Kunstmuseum in Bonn in 2001 (Figure 1). Museums and exhibitions as domains have already been explored in several research projects. Many of them focus on the goal to provide content concerning the artworks [9], or to immerse the user in a virtual augmented environment built in virtual museums [4], or to provide orientation and routing functionalities [5]. The visitors of the prototype experience personalized audio information about exhibits through their headphones. In July 2003, the first public exhibition is presented in the mentioned museum.

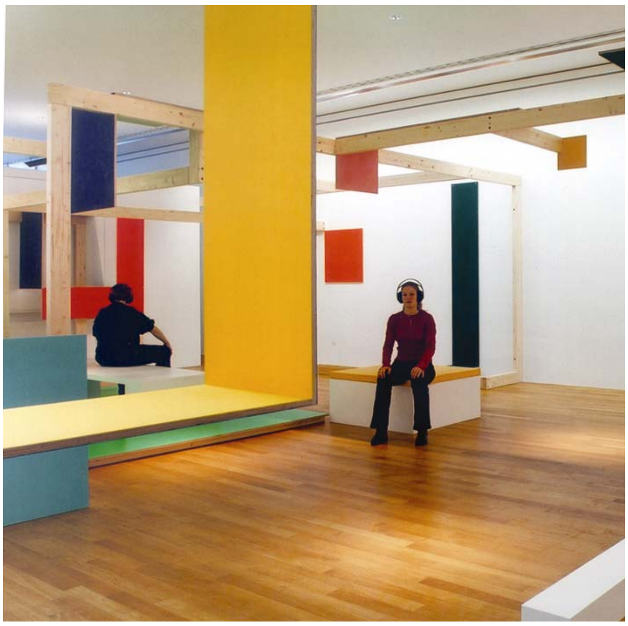

Fig. 1. Image of the LISTEN System Applied at the Kunstmuseum in Bonn

In the following sections we focus on the goals to be achieved, the modelling of an audio-augmented environment, and on the possibilities of combining the order, the source, and the content of sound items in order to retrieve a tailored presentation and provide a customized user experience.

3.1 Modelling the Audio-Augmented Environment

Combining high definition spatial audio rendering technology with advanced user modelling methods enables context-sensitive audio-augmented environments. Speech, music and sound effects are dynamically arranged to form a personalized and situated soundscape, offering information related to visual objects placed in the scenery as well as creating context-specific atmospheres. In order to seamlessly integrate the virtual scene with the real one a sophisticated auditory rendering process has to take into account the information provided by a very precise user tracking system. This tracking system observes the user's movements and delivers his/her current position and head orientation to the several modules of the LISTEN system (Figure 2).

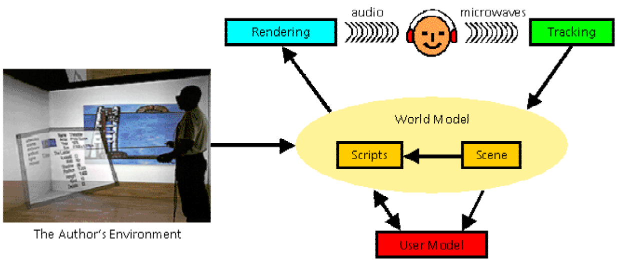

Fig. 2. Modules of the LISTEN System

To present customized information and create augmented environments, such systems have several models in common [10], described in more detail in this section: the world model, the augmentation layer, the domain model and the user model.
The *world model* (*space model*, *location model*) describes the physical environment the user moves through interacting with the system. In the LISTEN environment, the space model contains the geometric information of the environment and its visual objects. The LISTEN world model is a detailed virtual reality based geometric model. It is described as a geometric scene graph created by the AVANGO application [19]. The AVANGO system is a distributed virtual reality framework, which is used for authoring the physical domain in LISTEN, for example a museum.

The *augmentation layer,* on top of the world model, defines areas (Zones, Segments, Triggers) within the world model. By defining zones and segments, the user's focus obtains a valuable meaning and thus, the system is able to determine connections between the user's physical position and any physical or virtual object that is attached to this location. In Figure 3 we show the scenery of the application of LISTEN in a museum environment. The zones, which are connected to the visual objects placed within the environment, are divided into *object zones* and *near fields*. The object zones are associated with specific sound information about their connected visual object from a general perspective. The near fields are connected to smaller parts of the visual objects and contain more detailed sound information. The left hand picture shows an example of a sound emerging directly from the painting. The LISTEN system also allows the sound sources to be placed to certain parts of the painting or to be moving around.

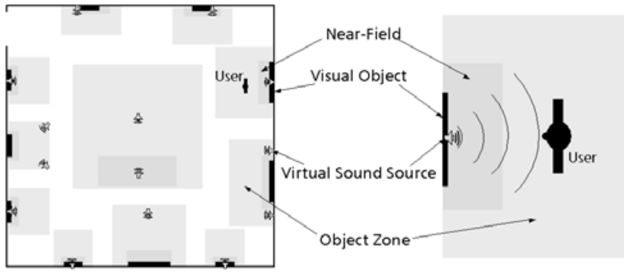

Fig. 3. Space Model of the LISTEN Application in a Museums Environment

The *domain model* holds information about sound objects and other hypermedia objects by using meta-data. The domain model builds up a virtual acoustic space, in which the location of virtual sound sources and spaces are defined. Stopping in front of a visual object generates aural information about the art piece. Moving the head and body activates a further audio source, where music deepens the user's impressions, or the voice of a commentator talks about the artist or describes the period the painting originates from.

The *user model* contains knowledge and profile information about the system's users. From the user modelling point of view, the AVANGO application serves as the target application, in which all the adaptation takes place While the user moves in physical space, AVANGO sends events to the personalization engine, which refines the user model. Furthermore, adaptation procedures are executed by AVANGO, in order to perform changes within the environment. A detailed description of the personalization process can be found in the next section.

4 Context-Awareness in LISTEN

Nowadays, adaptive systems mostly adapt information selection and presentation to the user's goals, preferences, knowledge, and interests. In most cases, the user model acquisition is driven by monitoring the activities of users concerning the user interface or by an analysis of specific device characteristics. In the HIPS project [14] the user's preferences, knowledge, interests, and movements in physical space are modelled to achieve a better fit of information and the users current context. Additionally, nomadic information systems make use of localization technologies to adapt to a richer context model. Besides the identification of the user's current location, appropriate technologies allow a precise user tracking in physical space.

4.1 Modelling the Context

One of the main objectives of the LISTEN project is the emphatic abandonment of any portable device for the user except the headphones. For the personalization process of LISTEN this means no explicit feedback from the user, which is would be an advantage for generating optimal recommendations. Since explicit feedback is not available in our application, the personalization process of LISTEN is based on the dimensions time, position and head orientation, and on an extensive amount of annotations for sound pieces amount of annotations for sound pieces respectively visual items.

The enrichment of information items with significant meta-information enables the personalization and customisation of information offers. We designed a domain ontology for the Macke exhibition to enable the description of these information items, thus allowing an individualized sequencing and presentation of these items. Paintings and sound items can be classified into this category system with a variety of dimensions describing the sound items technically and from a stylistic point of view. The ontology provides classification based on:

- technical descriptions of the sound items such as length of item, type (music, speech, sound effects);
- the relation to the physical space objects (art objects to which an item contributes, physical area zones or focuses to which they are connected);
- phases of work, image genre, or art technical aspects;
- the preferred target group, like the stereotypical listeners for such a sound item, or the emotional impacts or dramaturgy.

In particular speech sound items could be further classified into subcategories like citation, collage, diary, letter, newspaper and others to describe their style of presentation.

The LISTEN user modelling component can build hypotheses about the user's interests and preferences based on the time spent listening to certain audio streams or on the time spend gazing at a certain part of a visual object.

In our approach the personalization process is divided into four steps: Information collection, modelling, controlling and rendering. Each step fulfils a certain role within the user modelling process. Based on these four steps of the personalization process, the LISTEN application is able to adapt the scenery in various ways. The next subsections describe these modules in more detail.

4.1.1 Information Collection

The utilization of implicit feedback is an important issue for the realization of a personalized immersive environment. The only information to base the personalization process on is the spatial position of the visitor, the orientation of the head and some meta-information about sound snippets and visual objects. This meta-description consists of key-value pairs that hold information about colours, genre, epoch, etc. (e.g. colour = "blue"). A network of sensors is placed in the environment and connected to variable parameters of the domain. These sensors are used for recognizing changes within the environment and especially for the perception of the users' interaction with this environment. An observation module receives all incoming events 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, which can be further processed.

4.1.2 Modelling

By the means of statistical models, the implicit user profile already allows the deduction of valuable information that can be used for standard adaptation activity (e.g. the more time the user spends with a specific visual object, the more s/he likes it). This deducted information builds a behaviour model of the user. The behaviour model can be treated like an explicit representation of an implicit user feedback and may be consulted to draw an assumption about the user's interest in a specific object. Different machine learning and data mining algorithms are implemented to extract semantically enriched information and to gain more significant information relating to the user's behaviour.

In our user modelling approach for a museum environment we chose to employ an interest model and stereotypes not explicitly specified by the user. The interpretation of the user's focus in association with the time dimension helps to derive the users interests in the observed objects. By the interpretation of the users spatial position and head orientation based on the world model and augmentation layer the user's focus of attention can easily be derived. This focus points to a visual object, which is in turn associated with a meta-description. In combination with time the user spends looking on this visual object, the interest model is continuously refined with the corresponding meta-description. Starting with an empty representation, this accumulation procedure generically builds up a 2-depth tree. The nodes on the first tree level are named after the keys of all meta-descriptions that have been processed so far. The children of these nodes represent the accordant key-values and have a numeric value assigned that expresses the strength of the user's interest in this topic. This numeric value is increased every time a meta-description with the appropriate key-value pair appears as an input. To fine-tune the user's interest model, the LISTEN system processes the meta-description of the user's object of attention in a constant three-second interval. In succeeding steps this interest model supports the recommendation, filtering and collaboration process. Since this model is implemented in a domain independent way, it may be employed in several domains containing objects described by meta-information.

The use of stereotypes is a common approach in adaptive systems. With the aid of stereotypes, the personalization engine defines the users' observation type, thus the system is able to accordingly adapt the scenery. Every stereotype causes a different presentation and different strategies (see next subsection). In a museum environment it is not trivial to predefine meaningful stereotypes. Some stereotypes, which are easy to identify, are for example "adults" and "children". The stereotypes used by the LISTEN system are motion styles that represent stereotypical user behaviour in moving through an exhibition. In our domain the user may be *sauntering around*, *goal-driven* or *standing still* being either *focused* on a certain object or *unfocused*. This motion styles emerge from the analysis of the location and focus of the user and their evolution over time (history). These stereotypes result from the first review process of the LISTEN system and are meaningful, easy to detect and to revise. The interpretation of the user's motion style in combination with the location determinates the presentation style and facilitates the pre-filtering process of relevant sound pieces. If the user stands still focusing the object, object-dependent information is presented. If the visitor moves slowly with his/her look not being focused on one specific object, a zone-dependent, more general presentation starts.

4.1.3 Controlling

Meaningful user profiles accurately document the users' activity within the environment and can be exploited to adapt this environment in order to support the user and to provide personalized information. Therefore, 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. On the one hand, the control layer must consult the user model to get and set information about individual users, groups of users, or average behaviours in a certain environment. On the other hand this layer must be aware of the domain model to get information about the environment the user is currently moving in.

Based on these information sources the control component assembles a sequence of commands in order to adjust certain variable properties of the environment. Thus, different sequences of commands lead to different kinds of information presentation. This adjustment of environmental parameters is used for realizing domain independent adaptive or strategic methods, which are expressed domain dependently. On the strategic level of the control layer a decision between high level adaptations and the selection of different adaptive methods is taken.

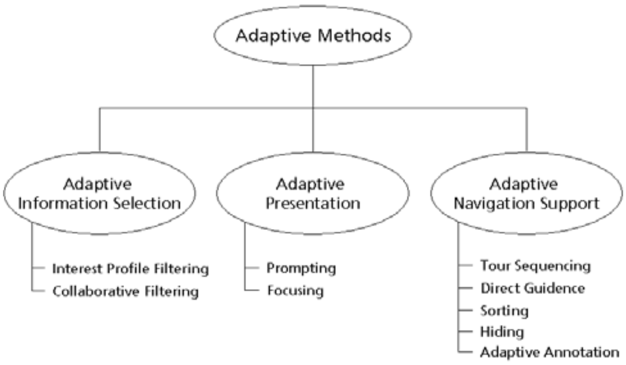

Fig. 3. Possible Adaptive Methods in Context-Aware Systems

In the field of adaptive hypermedia, several methods of adaptation and adaptive strategies are defined (cf. Figure 3) that are suited to tailor information selection and presentation to the individual user. To adapt these methods to the LISTEN principle, we have analysed the expressive means mainly used in audio environments and selected some of the mentioned adaptive methods that seemed appropriate for the personalization of an audio environment. Every adaptive method realizes its adaptive activity by queuing a sequence of commands for the rendering engine of the application (i.e. the AVANGO system). A more detailed description of the adaptive strategies implemented in the LISTEN system can be found in the next subsection.

4.1.4 Rendering

Rendering means handling the connection back to the domain. This engine translates the assembled sequence of domain-independent commands into domain-dependent commands. The implemented domain-dependent methods directly change variable parameters of the domain (i.e. change the audio augmentation) according to the users' behaviour. Thus, the decisions taken by the controlling component are to be mapped to real world actions.

4.2 Adaptation Strategies

In LISTEN, several domain independent adaptive methods and strategies are defined that are suited to tailor the information selection and presentation to the individual user. A control component selects and instantiates a strategy for presenting objects in the user's environment. Examples for domain independent adaptation methods on a strategic level are "adaptive prompting" and "adaptive annotation" for objects. In order to map the above-mentioned methods to the LISTEN principle, we have analysed expressive means mainly used in audio environments and selected some that seemed appropriate for the personalization of an audio environment.

4.2.1 Adaptation of the Sound Presentation

The basis for every kind of adaptation in LISTEN is the presentation of the sound space. Besides the decision, which sound item is to be played, the way the sound presentation can be modified is 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 already accomplished.

4.2.2 Adaptation of the Space Model

In LISTEN a virtual zone holds several zone specific sound pieces and is mostly associated to one specific domain object. Because users enter and leave these zones in virtual space (AVANGO), they should adapt to the user's particular behaviour. 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 expand up to a predefined point (zone breathing), so that s/he is able to listen to the sound further on. In addition, a sound item that is not completed should be transportable from one zone to another.

4.2.3 Adaptation to Social Context

If users are spatially and temporal similar (e.g. a family walking through an audio-augmented museum), they might want to receive the same audio information. Through building such clusters of people, for example, a subsequent discussion about seen objects is possible. Vice versa, the breaking up clusters of people is also possible. This would lead to a better distribution of people among several objects. Of course, one precondition for this expressive method is the detection of people clusters.

4.2.4 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 level of interest corresponds to the complexity, the amount, and the style of already received information about one object and is transferred to succeeding objects. If one of these objects complies with the user's interests, the sound presentation directly steps in the right level of interest and sound items are played that are classified on the adequate information depth and style.

4.2.5 Adaptation to Movement and Reception Styles

Several kinds of common behaviour can be identified with people walking through the environment (e.g. clockwise in museums). Attractor sounds emitted from different sound sources are used to draw the user's attention on certain objects of the environment. Thus, entire tours through the environment can be recommended. The selection and dynamic adaptation of tour recommendations can be adapted to the stereotypical type of movement and her/his preferred perception style. For examples of stereotypical user classes and their movements in museum environments see [21].

5 Evaluation

In order to refine the LISTEN system before its public presentation in July 2003 at the art exhibition in Bonn, some demonstrations already took place at the Kunstmuseum. The outcomes of these tests enabled a preliminary evaluation: in two workshops in which testing visitors, art curators, sound designers, and artists were involved, impressions and refinement issues were brought out.

A first positive result, reported by the 15 experts testing the system, was the success of the synesthetic experience: the visitors enjoyed the combination of audio-visual perception and felt as the interaction with the real visual objects was augmented. At the same time, curators appreciated the possibility to deliver content concerning the artworks in an innovative, enriched and less descriptive way.

Critical points in the domain model were also noticed: the zones of interaction surrounding each artwork were sometimes too small, thus forcing the visitor to approach the artwork very closely. In particular, in the case of overlapping zones, the boundaries of the object zones were hardly localizable by the user. In order to overcome these problems, we are working at making the zones more flexible, creating some "breathing zones" in which a sound is more attached to the user's behaviour in observing visual objects. In this sense, auditory icons providing some landmarks in the virtual environment navigation are meant to be inserted in the audio presentation in order to make the user aware of the interaction with the environment.

Further effort needs to be put into the recognition of the users' real focus as well: the tracking system senses the visitor's position in the space, but his focus can be on an object belonging to another sound presentation zone. Besides, some visitors 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.

We learnt that one of the main drawbacks of the LISTEN system would be *the lack of explicit user feedback* we have to cope with. Computer users are known for providing little feedback to rate the quality of items recommended to them. The lack of explicit feedback causes difficulties in clearly distinguishing between interesting and non-interesting objects. We can assume that in case of the LISTEN system, users will provide no (or at most very little) explicit feedback. In these situations, some systems use heuristics to determine positive and negative evidence of the user's information interest (i.e. unselected objects are negative examples [18]). Another approach developed in [13] uses significant analysis aiming at selecting those features that are extraordinarily important to the user for identifying relevant objects.

Another main criticism in the evaluation was the use of stereotypes (i.e. fact oriented, general overview oriented, or emotional) that were to be selected in advance before the presentation started. The visitors do not like to be clustered and classified: the request to select what kind of art visitors they are generated irritation. Due to this fact and due to the decision of not providing any input devices (neither desktops nor handhelds) we intend not to use stereotypes in refining the system, but rather gather more significant information about the user. The new test scenario will provide hidden stereotypes that represent the user's moving styles like *sauntering*, *goal-driven* or *standing still*, because these stereotypes are meaningful, easy to detect and to revise. We have to approach the fact that in the LISTEN system the *human-computer interaction only lasts a short time period*. Therefore, the system needs to pre-process a sufficient amount of past users' data in order to build a meaningful user model. Since such adequate information opulence is not accessible by now, we are forced to develop learning algorithms that can infer from a small set of data. The next evaluation step will provide much more information, when the LISTEN system was experienced by the public.

6 Conclusions

In this paper we presented an approach for adapting the sound presentation in an audio augmented environment. The technology for this sound augmentation is provided by the LISTEN system, into which a user modelling component was embedded. This component assembles a bunch of well-known user modelling techniques for adapting the audio presentation regarding to the users' preferences, interests, and motion. We discussed the application of these techniques in our special environment and described suitable realizations.

We attempt to highlight the potential of intelligent audio user interfaces, 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 user and context modelling process of an intelligent audio-augmented environment.

In the LISTEN application, the only user interface is the user's motion. This implies a trade-off in terms of communication between the user and the system. Once the user has started his/her walk through the environment, s/he does not carry any device except the headphones. The only events that the user can send to the system are his/her positions in the time, but no explicit control channel is provided. Thus, the system will not get any explicit feedback from the user during the user's tour. Thus, the user can't switch to a different type of sound presentation while experiencing the environment, nor controlling how to play the sound presentation. To be able to provide audio sequences that suit to the user's context, the system needs to infer information from the user's behaviour. The basis for this inference process is the data delivered by the ence process is the data delivered by the observation component:

- Time (e.g. timestamps, delays, time sequences),
- Position (e.g. position in the physical environment, position in the virtual environment like zone identifiers),
- Focus (e.g. the identifier of an object the user is looking at).

In addition to the adjustment of the audio presentation associated with a specific object, a certain strategy decides which particular object is recommended next. All intended adaptation processes result in several sequences of commands for controlling the environment.

References

1. Balabanovic, M. and Shoham, Y. F., (1997), Content-Based, Collaborative Recommendation. In: Communications of the ACM, March 1997, Vol.40, No. 3, pp. 77-87.
2. Basu, C., Hirsh, H. and Cohen, W. (1998), Recommendation as Classification: Using Social and Content-Based Information in Recommendation. In: Proceedings of the Fifteenth National Conference on Artificial Intelligence, Madison, WI, pp. 714-720.
3. Billsus, D., and Pazzani, M. J. (1999), A Hybrid User Model for News Classification. In: Kay J. (ed.), UM99 - User Modeling Proceedings of the Seventh International Conference, Springer-Verlag, Wien, New York 1999, pp. 99-108.
4. Chitarro, L., Ranona, R., Ieronutti, L. (2003): Guiding Visitors of Web3-dimensional Worlds through Automatically Generated Tours. In: Proceedings of Web3D 2003: 8th International Conference on 3D Web Technology, to appear.
5. Ciavarella, C., Paternò, F.: Supporting Access to Museum Information for Mobile Visitors. In: 10th International Conference on Human-Computer Interaction, Crete, Greece, 2003
6. Delgado J., Ishii N., Ura T. (1998), Content-based Collaborative Information Filtering: Actively Learning to Classify and Recommend Documents. In: Proceedings of the CIA'98, Paris, France, July 1998, LNCS 1435. Springer Verlag, Wien, New York 1998.
9. Dey, A.K. & Abowd, G.D. Towards a Better Understanding of Context and Context-Awareness. GIT-GVU-99-22, College of Computing, Georgia Institute of Technology, ftp://ftp.cc.gatech.edu/pub/gvu/tr/1999/99-22.pdf, 1999.
8. Eckel, G. (2001), 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.
9. Gaver, W. (1994): Using and Creating Auditory Icons. In: Kramer, G. : Auditory Display: Sonification, Audification and Auditroy Interfaces. Addison Wesley, Reading MA.
10. Goßmann, J. & Specht, M. (2001). Location Models for Augmented Environments. In the Workshop Proceedings of *Location Modelling for Ubiquitous Computing*, Ubicomp, pp. 94-99.
11. Gross, T., Specht, M.. *Awareness in Context-Aware Information Systems*. In Mensch & Computer – 1. Fachübergreifende Konferenz (March 5-8, Bad Honnef, Germany). H. Oberquelle, R. Oppermann and J. Krause, Eds. Teubner, 2001. pp. 173-182.
12. Maybury, M.T. (2002): User Modeling for Adaptive User Interfaces. Tutorial Notes of International Conference on Advanced Visual Interfaces (AVI2002), Trento.
13. Mladenic D. Personal WebWatcher (1996), Implementation and Design. Technical Report IJS-DP-7472, Department of Intelligent Systems, J. Stefan Institute, Slovenia, 1996.
14. Oppermann, R., Specht, M. (2000): A Context-sensitive Nomadic Information System as an Exhibition Guide. Proceedings of the Handheld and Ubiquitous Computing Second International Symposium, HUC 2000, Bristol, UK, September 25-27, 2000, 127 – 142
15. Pazzani, M. (1999), A Framework for Collaborative, Content-Based and Demographic Filtering. In: Artificial Intelligence Review 13(5), pp. 393-408.
16. Resnik P. and Varian H. (1997), Recommender Systems. In: Communications of the ACM, March 1997, Vol.40, No. 3, pp. 56-58.
17. Schilit, B.N., Adams, N.I. and Want, R. (1994), Context-Aware Computing Applications. In: Proceedings of the Workshop on Mobile Computing Systems and Applications (IEEE Computer Society, Santa Cruz, CA), pp. 85-90
18. Schwab, I., Pohl, W. and Koychev, I. (2000), Learning to Recommend from Positive Evidence, In: Proceedings of Intelligent User Interfaces, ACM Press 2000, pp.241-247.
19. Tramberend, H. (1999), Avango: A Distributed Virtual Reality Framework. In: Proceedings of the IEEE Virtual Reality '99 Conference, Houston,Texas, USA.
20. Unnützer, P. (2001), LISTEN im Kunstmuseum Bonn, KUNSTFORUM International, Vol. 155, pp. 469-70.
21. Zimmermann, A., Lorenz, A. and Specht, M. (2002), Reasoning From Contexts. In: Proceedings of the 10th GI-Workshop Adaptivity and User Modeling in Interactive Systems, Hannover, Germany, 2002, pp. 114-120
22. Zimmermann, A., Lorenz, A. and Specht, M. (2003), The Use of an Information Brokering Tool in an Electronic Museum Environment. Proceedings of the Museums and the Web, Charlotte, North Carolina, 2003