Inferring Interests from User Movements: The LISTEN-approach

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

Abstract

The paper describes our experience in adapting audio presentations in a museum environment without falling back on explicit user feedback. Our contribution introduces the LISTEN project conducted by the Fraunhofer Institut in St. Augustin and partners, discusses results of the evaluation process and illustrates our approach of exploiting the user’s movement as the only interface for interaction.

The LISTEN project deals with the audio augmentation of real and virtual environments. The system users move in physical space wearing wireless headphones and listening to a contextualized audio presentation. A realistic immersion of the user into the audio-augmented environment demands for the combination of both headphones, which are able to render 3d sound, and fine granular tracking information for the determination of the user’s exact position and orientation of the head. A first LISTEN prototype is applied to an art exhibition being launched at the Kunstmuseum in Bonn in July 2003. Visitors of this exhibition experience personalized audio information about art objects through their headphones.

Since the objectives of the LISTEN project emphatically disclaim the request of explicit user feedback, the utilization of implicit feedback is an important issue for the realization of a personalized immersive environment. Our contribution clarifies step by step how basic information like the user’s spatial position and head orientation, in combination with meta information for physical objects and sound snippets, are evaluated to conclude the user’s interest in the objects of the environment.

The definition of an overlay model within the system allows a segmentation of the physical space into virtual zones, and the connection of these zones to specific physical objects placed within the environment. The interpretation of the user’s physical movements in the context of this overlay model leads to valuable information about the user’s location (i.e. the entered zone), and the focus (i.e. the object catching the user’s attention).

During the evaluation process we have presented the implemented system to a variety of experts made up of museum curators, artists, and composers, who have appreciated the success of the synaesthetic experience: the visitors enjoyed the combination of audio-visual perception and felt as the interaction with the real visual objects was augmented. Namely, 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 noticed: the zones of interaction surrounding each artwork were sometimes too small, thus forcing the visitor to approach the artwork very
closely. 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. Especially 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 behavior 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. To define a more sufficient model of the user’s interest, his/her focus of attention and time factors have to be taken into account. Therefore, we infer stereotypes representing moving styles like sauntering, goal-driven or standing still (either focused on a certain object or unfocussed). The combination of the user’s motion style and location determinates the presentation style and facilitates the 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.

Finally, the selection of one specific sound piece to be played depends on the user’s history of already known sounds and an interest model. This interest model is based on the time the user spends with a certain object observing it and allows a ranking of sound pieces with regard to their relevance for the user. Our paper will explain the mechanism for the assignment of a relevance value to each sound piece in more detail. Additionally, the architecture of the implemented system ready to use will be specified.