# Tangible User Interface for the Exploration of Auditory City Maps

Martin Pielot[1], Niels Henze[1], Wilko Heuten[1], and Susanne Boll[2]

[1] OFFIS Institute for Information Technology, Escherweg 2, 26121 Oldenburg, Germany
{Pielot,Henze,Heuten}@offis.de
[2] University of Oldenburg, Escherweg 2, 26121 Oldenburg, Germany
Susanne.boll@informatik.uni-oldenburg.de

**Abstract.** Before venturing out into unfamiliar areas, most people scope out a map. But for the blind or visually impaired traditional maps are not accessible. In our previous work, we developed the "Auditory Map" which conveys the location of geographic objects through spatial sonification. Users perceive these objects through the virtual listener's ears walking through the presented area. Evaluating our system we observed that the participants had difficulties perceiving the directions of geographic objects accurately. To improve the localization we introduce rotation to the Auditory Map. Rotation is difficult to achieve with traditional input devices such as a mouse or a digitizer tablet. This paper describes a tangible user interface which allows rotating the virtual listener using physical representations of the map and the virtual listener. First evaluation results show that our interaction technique is a promising approach to improve the construction of cognitive maps for visually impaired people.

**Keywords:** sonification, auditory display, tangible user interface, spatial audio, exploration, interaction techniques, visually impaired users.

## 1 Introduction

The ability to travel is an important factor for social inclusion. Traveling to unknown areas relies on external information resources like city maps, which can be utilized at home during the travel preparation. Blind people are mostly not able to access these representations. Consequently, they usually do not travel to unknown areas and are excluded from many social activities. We developed "Auditory Map" which conveys geographic information to blind and visually impaired people through an auditory display [6, 7]. We identified the geographic entities that are most important to gain an overview of an area like lakes, parks, and buildings. They are represented by continuously playing corresponding natural sounds such as dabbling water or singing birds. These sounds are placed on a plane within a virtual sound room according to their locations on the map thus maintaining their spatial relations. Users walk through the room virtually by moving a virtual listener. The sound objects transmit their sound up to a maximum distance, so that the user perceives the sounds of the objects in the near environment of the virtual listener only. A lake on the left of the virtual listener is heard from the left and a lake on the right is heard from the right. A digitizer tablet

I. Oakley and S. Brewster (Eds.): HAID 2007, LNCS 4813, pp. 86–97, 2007.

serves as input device for the exploration. The city map is mapped on the tablet, which thus represents an absolute frame of reference. Moving the stylus over the tablet updates the virtual listener's position accordingly (see Figure 1).


**Fig. 1.** A user exploring an area with the Auditory Map

Our previous evaluations [6, 7] of the Auditory Map with blind and sighted participants showed that the exploration of auditory maps leads to survey knowledge about the presented area and to the ability of making spatial decisions. Nevertheless, we identified difficulties localizing auditory objects that occur in particular when objects are in front or behind of the virtual listener. This can be solved by enabling the user to rotate the virtual listener within the virtual sound room. However, the currently used digitizer tablet cannot serve as input device for rotating.

In this paper we therefore describe a newly developed tangible user interface for the exploration of virtual sound rooms. It combines the advantages of a digitizer tablet and its stylus with the possibility to rotate the virtual listener continuously. We assume that this interface will improve the access to the sound rooms' information for blind and visually impaired people, leading to more accurate mental representations of spatial information and providing more confidence in making spatial decisions.

In the following section, we analyze the causes of localization difficulties and ways to address them. Section 3 discusses the related work. Section 4 proposes the new tangible user interface for exploring virtual sound rooms followed by the description of its implementation using computer vision technology in Section 5. The preliminary evaluation of the tangible user interface is presented in Section 6. We close this paper with a summary and an outlook.

## 2 Reasons for Localization Inaccuracy

The Auditory Map uses an auditory display to convey geographic information. Auditory objects are arranged in a planar sound room, rendered by a 3D sound library, and conveyed using headphones. To build a realistic mental map of the presented area it is important to localize the geographic objects accurately. Localization accuracy is affected by certain aspects that we will highlight in this section. We also outline our concept to improve the localization accuracy by introducing rotation to the Auditory Map.

Evaluating our Auditory Map [6, 7] we observed that our participants had difficulties to perceive the directions of geographic objects accurately. We found three reasons for that:

1. Localization accuracy of sound sources is limited even in the real world due to limitations of the human perception [1] and in particular in virtual sound rooms [4] due to limitations of the acoustic rendering. This can be countered by using personalized spatial audio displays but is not feasible for applications that address a broader audience.
2. The sounds used to display geographic objects are important for localization accuracy. In Auditory Maps, geographic objects are represented by natural sounds. This limits the number of possible choices for the representing sounds. Thus sounds might have to be chosen which are not ideal from the perspective of localization accuracy.
3. The lack of context information in virtual sound rooms compared to reality leads to more front-back confusions [18]. In reality we move and rotate our head while listening to a sound source. We are thus able to combine the knowledge about the head's movement with the changing acoustic impression and construct a more accurate spatial model of the sound sources' locations.

To improve the localization inaccuracy we address this lack of context information. Minnaar et al. have shown that head movements can reduce the localization blur when listening to synthesized spatial sound [12]. Head movement also helps avoiding front-back confusion [10]. We therefore will enable the user to rotate the virtual listener continuously as if it would turn its head left and right. If the user is not sure whether an auditory object is behind or in front, he or she can rotate the listener and follow the movement of the sound source. The user can resolve front-back confusions in a self-directed way. However, the digitizer tablet is not suitable to alter the orientation of the listener. Thus, the interaction with the Auditory Map has to be revised. The following section will review existing work on which we based our design decisions.

## 3 Related Work

Sleuth [5] and AudioDoom [9] use a virtual listener to explore spatial audio information. The interaction is realized by a mouse and a ringmouse. Both input devices provide no frame of reference. To avoid the possible loss of orientation the freedom of interaction is restricted to 90° steps. Since we aim at continuous rotation, this solution is not appropriate for the Auditory Map. The Nomadic Radio [14] by Sawhney and Schmandt utilizes speech recognition to manage voice and text-based messages with an auditory display. Cohen [3] proposed a gesture-based interface, where users could point, catch and drag spatially arranged sounds. Both methods are not suitable for our task because they do not provide feedback about the virtual listener's position and orientation.

Tangible user interfaces enable users to interact with digital information systems over tangible physical artifacts. Each physical artifact represents and allows manipulation of the digital information it is associated with [16]. Thus, the physical artifacts provide haptic feedback about the associated digital states. The ActiveCube of

Watanabe et al. [17] has shown that tangible user interfaces are well suited to convey spatial information to visually impaired and blind people. The project Tangible Newspaper [15] shows, that tangible user interfaces are also well suited for obtaining an overview about complex documents such as maps. Boverman et al. [2] did promising research at exploring spatially sonified information with tangible user interfaces. They allow users to scan the information by “exciting” them with a physical artifact. These projects indicate that tangible user interfaces are well applicable for the exploration of the sound room used in the Auditory Map.

## 4 A Tangible User Interface for Exploring Virtual Sound Rooms

Exploiting the previously discussed advantages, we developed a novel user interface to integrate rotation into the Auditory Map. We did not use the user’s head movement to rotate the listener of a virtual acoustic room since that would divide the control of position and orientation into two modalities. It would also require the user to perform 360° turns. However, as we use the position of the Auditory Map stylus input device to move the virtual listener we could also use the orientation of the input device to adjust the orientation of the virtual listener.

The interaction technique based on the digitizer tablet was highly appreciated by our users. We assume that the characteristics of this technique are a good guideline to develop an interaction that enables the user to move and rotate the virtual listener. In the following section we analyze the requirements for the enhanced interaction. Afterwards we determine the physical artifacts for the tangible user interface and design the user interaction.

### 4.1 Requirements

When exploring a map, position and orientation of the virtual listener are not displayed explicitly by the auditory display. This lack of feedback can lead to the loss of orientation as described by Ohuchi et al. [13]. It would also become more likely that the user cannot distinguish between self-motion and movement of the perceived objects [8]. When using the digitizer tablet to interact with the Auditory Map the tablet serves as a frame of reference since the relation between the surface and the stylus translates directly to the position of the virtual listener in relation to the border of the presented map. This allows the user to haptically identify the virtual listener’s current position. For introducing rotation a similar frame of reference, haptic feedback, and intuitive mapping is needed. We identified the following five requirements. The tangible user interface has to:

- enable the user to place the virtual listener at any location on the displayed map,
- enable the user to continuously rotate the virtual listener around the vertical axis,
- provide feedback about the virtual listener’s position,
- provide feedback about the virtual listener’s orientation, and
- provide this feedback immediately.

### 4.2 Determining the Artifacts

Regarding these requirements we concluded that the map and the virtual listener need to be represented as physical artifacts. To determine appropriate physical artifacts Ullmer and Ishii describe three common approaches [16]: Their least favorable approach subordinates aesthetical aspects to technical issues. The artifact is chosen by its functionality and not because it suites well from the perspective of usability and design. Their most favorable approach is the augmentation of existing physical artifacts while retaining their typical usage. This allows people to apply their existing knowledge to the new interface. Their third approach uses so called “found objects” and is favorable if the previous approach is not applicable. If the right objects are chosen, the user can suggest their functionality by their appearance.

Choosing the physical representation for the map we followed the second approach. We took a tactile map as archetype and identified the paper on which the map is printed as the complement to the map of the auditory displayed map. Instead of a paper sheet we chose a felt pad as the physical representation of the map because its borders are easier to perceive haptically. The size of the pad shown in Figure 2 is about 16x12 inches (40x30 cm). We marked the northern border with an adhesive strip to make the cardinal directions identifiable by touch.

Since the virtual listener is a unique feature of the Auditory Map, the choice of a physical artifact is not as straight-forward as for the map. Common tactile maps are explored by the user’s finger. Following the preferred approach the user’s finger would serve as representation for the virtual listener. But when exploring tactile maps the orientation of the finger does not alter the perception of the map in the way rotation would alter the perception of the auditory display. In addition, using the finger would require uneconomical gestures of a user. Thus we decided against using the user’s finger for the interaction. Instead we used a “found object”. For the prototype we chose a toy duck, which is shown in Figure 2. Due to the characteristic physique of ducks, visually impaired users can determine its orientation by touching it.

**Fig. 2.** Exploring the Auditory Map by moving and rotating the virtual listener’s artifact

Users explore the Auditory Map by moving and rotating the duck while it is located on top of the pad. The position and orientation of both artifacts in relation to each other is transferred immediately to the virtual listener's position and orientation. Users can rotate the duck and the pad to change the orientation of the virtual listener. For example, if the virtual listener’s artifact is located in the south western corner of the pad and faces the eastern border, the virtual listener is placed in the south western corner of the map, and turned eastwards. If the user moves the virtual listener's artifact to the middle of the pad and turns it north, the virtual listener does the same movement (see Figure 2). This enables the user to determine position and orientation of the virtual listener inside the map by examining the state of the tangible user interface.

## 5 Implementation

The concept described in the previous section has been realized by implementing the tangible user interface and integrating it with the existing Auditory Map. In order to make the tangible user interface work, it has to track the position and orientation of the physical artifacts that represents the map and the virtual listener. Tracking the artifacts must be performed with short latency. In order to build a reliable and low-cost system we decided to use computer vision technology based on inexpensive webcams. We implemented the visual object tracking using and extending the Intel Open Source Computer Vision Library (OpenCV). The artifacts' relative positions and orientations are used to calculate the virtual listener's position and orientation inside the map. We assume that the webcam is mounted perpendicular to the pad artifact and the virtual listener's artifact is on top of the pad.

### 5.1 Follow the Map Artifact

To track the position and orientation of the pad we implemented a three-stage process that takes place after calibrating the algorithm. First, we separate the artifact from the background, then we identify the artifact using previous knowledge about its size and shape, and finally we determine its position and orientation. Figure 3 shows the image of the webcam on the left and the recognized position of the map's artifact after background subtraction in the centre.

Prior to using the tangible user interface the setup has to be prepared by putting the pad into the area captured by the webcam. The system recognizes the pad by its rectangular shape and assumes that the artifact's current top edge symbolize the north-edge of the map. To separate the artifact from the background we implemented a non-adaptive background subtraction [11]. This technique compares a static prepared image of the background pixel-wise with current images of the camera. If the difference between a pair of pixels is greater than a certain threshold the pixel is considered as a part of foreground objects. The other pixels are considered to be background and turned black. In the resulting foreground image the algorithm looks for edges that form connected graphs. Graphs without at least one pair of orthogonal edges are discarded. Pairs of orthogonal edges are being used trying to approximate the pose of the artifact by taking the artifact's position detected in the previous video frame into account. Finally, the cardinal directions of the edges are determined by comparing the artifact's position and orientation with the previous position and orientation. The algorithm is robust against occlusion as long as one edge is completely visible and one of its adjacent edges one is visible.

### 5.2 Track the Virtual Listener's Artifact

While users interact with the tangible user interface the listener's artifact will always be in front of the map's artifact. Thus, it can not be identified as a single object using background subtraction. Since we do not want to restrict the choice of the artifact for the virtual listener, we also cannot utilize knowledge about its shape as we did with the map. Hence, we choose a tracking method that is independent from background subtraction and the artifact's shape. The chosen method relies on markers to track the artifact's position and orientation. The artifact has to be tagged with two specific markers as shown in the right of Figure 3. One marker is fixed at the front and the other at the back of the artifact. The markers are identified by the hue with must be unique. The algorithm searches for pixels with each marker's hue. Afterwards, the center of the found pixels is calculated. By using the vector between the centers of each centers, as indicated by the line in the right of Figure 3, the position and orientation of the artifact is determined. Once the system is calibrated to the markers' hue, the detection runs stable and recovers quickly from distractions like temporary occlusion of the artifact.

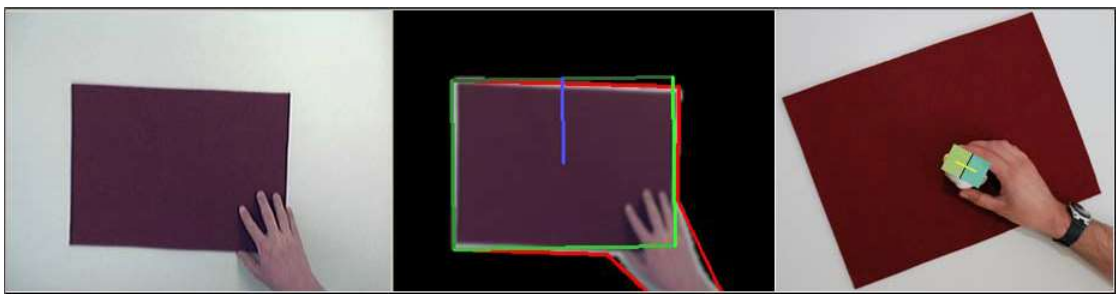

**Fig. 3.** Tracking the map's artifact (left and centre) and marker-based detection of the virtual listener's position and orientation (right)

## 6 Evaluation

We conducted a preliminary evaluation with the implemented tangible user interface. We gathered quantitative and qualitative data by conducting three experiments and afterwards handing out questionnaires to the participants. The goal was to get indications about whether rotation and the tangible user interface are useful extension to the Auditory Map.

### 6.1 Evaluation Setup and Methods

The evaluation was performed by eight sighted participants. All of them were between 16 and 35 years old. No one had previous experience with the Auditory Map or the tangible user interface. All participants were sighted but blindfolded prior to the evaluation. Thus they had to make themselves familiar with the system without vision. We assume that sighted blindfolded users are well suited to provide preliminary feedback about the tangible user interface. Later evaluations will be conducted with the target user group.

Figure 4 shows the system setup that was used during the evaluation. The left image shows the setup of the webcam that was mounted on a holder perpendicular to a table below. The center image shows a participant interacting with the tangible user interface. The right image shows the map used in the second and the third experiment. The lake is represented by dabbling water, the park by singing birds, the church by the sound of church bells, and the visitor attraction by the noise of several single-lens reflex cameras. The dashed lines indicate the route that should be followed in the one of the experiments.


**Fig. 4.** System setup (left), a user interacting with the tangible user interface (center), and scheme of the map used in the experiments (right)

We used protocol sheets, a photo camera and a video camera to log the experiments, their results and the participant's notes. The protocol sheets were used to write down the results of each experiment and the annotations made by the participants. The state of the tangible user interface during the experiments was recorded with the photo camera. The video camera was used to monitor the participant's interaction with the tangible user interface during the experiments for later analyses. Following the experiments we handed out questionnaires to the participants. They contained multiple choice questions where the participants could rate aspects of the tangible user interface. Further open questions encouraged the participants to write down their impressions, considerations and suggestions.

### 6.2 Experiments and Results

During the evaluation we conducted three experiments. The experiments should evaluate whether the tangible user interface acts as a frame of reference, whether rotation helps to increase the localization accuracy and whether rotation as an interaction technique is accepted by the users.

**Experiment 1.** The first experiment should show if the tangible user interface serves as a frame of reference for the position and orientation of the listener. It consisted of two tasks. The first task was to determine the listener's orientation, which was placed randomly on the map before. The second task was to place it exactly at the center of the map and align it exactly north. The experiment was conducted without the auditory output of the Auditory Map. Since the participants were blindfolded, they had to rely solely on their sense of touch.

The deviation of the listener's orientation the participants determined in the first task averaged 1.6°. In the second task, the deviation of the alignment with the north-south axis averaged 5.5°. The deviation of the listener's position from the map's center averaged 2.0% east-west and 4.3% north-south of the map artifact's size. The results show that users can determine the listener's position and orientation with only slight deviations. The experiment showed that it is easy for users to determine the position and the orientation of the listener through the haptic sense only when interacting with the tangible user interface. Thus the tangible user interface can serve as the frame of reference.

**Experiment 2.** The goal of the second experiment was to examine the impact of rotation on the localization of landmarks. It consisted of two tasks and took place on a map (see Figure 4) containing four landmarks. In the first task the listener was placed at the map's center, aligned north and the users were not allowed to alter its position and orientation. The participants then had to estimate the direction of three landmarks. In the second task, they were encouraged to rotate the listener without moving it And re-estimating the directions of the three landmarks. Comparing the directions determined during the two tasks should show whether rotation helps to improve the localization accuracy and helps to resolve front-back confusions.

The results of the second experiment differed significantly depending on the sound of the landmark, the users were asked to localize. The localization of the lake showed that rotation can help in overcoming uncertainties of localizing landmarks. Without rotation five of the eight participants were not certain of the lake's direction. With rotation every participant was able to estimate a direction. Nevertheless we still observed front-back confusions in two cases. Every participant was able to localize the church. The comparison of the average deviations with and without rotation showed that rotation improves the accuracy of determining it direction. The deviation of the directions given by the participants was reduced to 15.4° and a median of 10° with rotation, compared to 27.75° average and a median of 25° without rotation. The localization of the park showed the importance to choose well distinguishable sounds for landmarks when allowing rotation. All participants stated that they had problems to distinguish the park's bird chanting from the camera noise of the attraction. Without rotation allowed, two of the eight participants were uncertain about the park's location. With rotation allowed four participants could not locate the park. The results of the second experiment indicated that rotation improves the localization of landmarks only if their sound can be easily distinguished from the other landmarks' sounds.

**Experiment 3.** The third experiment should show whether rotation is accepted by the users and represents a valuable enhancement of the interaction with the Auditory Map. We wanted to observe, if users make use of rotation spontaneously. The participants were asked to follow and describe a route between four landmarks (see Figure 4). They were asked to determine the relative direction changes along from the perspective of a person that walks along this route. This task should implicitly encourage the participants to use rotation.

While exploring the route, the participants were allowed to interact with the interface without restrictions. The video camera was used to record the participants' interaction while accomplishing the task. The videos were used to analyze if and how rotation was used to follow and describe the route. All participants were all able to give the direction changes with little deviation. All participants rotated the listener spontaneously to localize subsequent landmarks and to determine the relative direction change. This indicates that rotation is well accepted and found valuable by the users.

**Overall user observations.** To get an idea about the participants' impression we analyzed the questionnaires and the participant's notes. Most participants found it easy to determine position and orientation of the listener in the first experiment. In the second experiment most of them found it hard to localize landmarks, especially without rotation. Although rotation only improved the localization of the church and the lake, most participants noted that rotation eased the localization of landmarks for them. They also found rotation useful for following the route in third experiment. Many participants stated that localizing the park and the lake was difficult, while localizing the attraction and the church was easy due to their respective sounds. Many participants found the tangible user interface easy to understand. One participant stated that the duck especially helped to determine the relative direction changes. Some participants found it difficult to concentrate on the task while listening to the auditory presentation of the map and proposed to make the volume adjustable. The participants only rotated by turning the listener's artifact, not the map's artifact. Most users localized landmarks by turning the listener's artifact until facing it.

### 6.3 Discussion

Since the participants found rotation useful and spontaneously made use of it during the third experiment, we assume that rotation is a useful enhancement of the interaction with the Auditory Map. The localization of the well distinguishable sounds improved with rotation while the localization of the badly distinguishable sounds did not. This indicates the importance to choose easily distinguishable sounds to represent landmarks. Rotation has also been used to resolve uncertainties and front-back confusions, though they were not always resolved correctly. We assume that the front-back confusions that occur with rotation were due to the participants' lack of experience.

Even though we observed uncertainties and front-back confusion with rotation, none of the participants lost orientation during the evaluation. This indicates that the tangible user interface is a well suited frame of reference. The participants had no problems to interact with the Auditory Map even though they had not seen the tangible user interface prior to the experiments. The participants found the tangible user interface easy to understand and showed no problems to interact with the Auditory Map.

## 7 Conclusions

In this paper we presented a novel tangible user interface for the exploration of virtual sound rooms by moving and rotating a virtual listener through them. We integrated it into our previous work, the Auditory Map, which utilizes spatial non-speech sound to present geographic entities and their spatial relations. The user interface consists of two physical artifacts representing the map and the virtual listener. Blind users now simply interact with the Auditory Map by moving and rotating physical objects. These manipulations are tracked by computer vision technology and are mapped accordingly to the virtual listener and map.

Our evaluation showed that the physical objects provide immediate feedback to the user about the virtual listener's location and orientation on the map. We observed that the participants intensively used rotation and that they found rotation very helpful to determine the direction of the geographic objects. We assume that the tangible user interface would be even more effective for trained users. The combination of presenting information through spatial sound and using tangible user interfaces to explore this information has great potential to provide access to any spatially distributed data. In the future further evaluations with the target user group will be conducted. Another aspect of our further research will be to advance the tangible user interface by integrating panning and zooming into auditory maps.

## References


1. Blauert, J.: Spatial Hearing: The psychophysics of human sound localization. MIT Press, Cambridge (1999)
2. Bovermann, T., Hermann, T., Ritter, H.: Tangible Data Scanning Sonification Model. In: Conference on Auditory Display (2006)
3. Cohen, M.: Throwing, pitching and catching sound: audio windowing models and modes. International Journal of Man-Machine Studies 39(2), 269–304 (1993)
4. Donker, H., Klante, P., Gorny, P.: The design of auditory user interfaces for blind users. In: Nordic conference on Human-computer interaction (2002)
5. Drewes, T., Mynatt, E., Gandy, M.: Sleuth: An Audio Experience. In: International Conference on Auditory Display (2000)
6. Heuten, W., Henze, N., Boll, S.: Interactive Exploration of City Maps with Auditory Torches. In: Computer-Human Interaction (2007)
7. Heuten, W., Wichmann, D., Boll, S.: Interactive 3D Sonification for the Exploration of City Maps. In: Proceedings of the 4th Nordic conference on Human-computer interaction (2006)
8. Larsson, P., Västfjäll, D., Kleiner, M.: Perception of Self-motion and Presence in Auditory Virtual Environments. In: Conference on Auditory Display (2004)
9. Lumbreras, M., Sánchez, J.: Interactive 3D sound hyperstories for blind children. In: Computer/Human Interaction Conference (1999)
10. Marentakis, G., Brewster, S.A: Gesture Interaction with Spatial Audio Displays: Effects of Target Size and Inter-Target Separation. In: Conference on Auditory Display (2005)
11. McIvor, A.M.: Background subtraction techniques. In: Proceedings of Image and Vision Computing (2000)
12. Minnaar, P., Olesen, S.K., Christensen, F., Møller, H.: The Importance of Head Movements for Binaural Room Synthesis. In: Conference on Auditory Display (2001)
13. Ohuchi, M., Iwaya, Y., Suzuki, Y., Munekata, T.: Cognitive-map Formation of Blind Persons in a Virtual Sound Environment. In: Conference on Auditory Display (2006)
14. Sawhney, N., Schmandt, C.: Nomadic radio: scaleable and contextual notification for wearable audio messaging. In: Computer-Human Interaction Conference (1999)
15. Sporka, A.J., Němec, V., Slavík, P.: Tangible newspaper for the visually impaired users. In: Computer/Human Interaction Conference (2005)

16. Ullmer, B., Ishii, H.: Emerging frameworks for tangible interfaces. IBM Systems Journal 39(3-4), 915–931 (2000)
17. Watanabe, R., Itoh, Y., Asai, M., Kitamura, Y., Kishino, F., Kikuchi, H.: The soul of ActiveCube: implementing a flexible, multimodal, three-dimensional spatial tangible interface. Computers in Entertainment 2(4), 15–15 (2004)
18. Wightman, F.L., Kistler, D.J.: Resolution of front-back ambiguity in spatial hearing by listener and source movement. Journal of the Acoustical Society of America 105(5), 2841–2853 (1999)