6 AUGMENTED AUDIO

Each thing we see hides something else we want to see. — René Magritte

AUGMENTED Audio Reality (AAR) describes a system that *enhances* the acoustics of a real world location through the display of additional auditory information, thus allowing others to see things as *we* want them to be seen. Augmented audio is an extension and a special case of the 3D virtual auditory environments that were discussed in the last chapter. The focus of this chapter is a detailed analysis of augmented audio reality in terms of acoustic presentation, listener immersion and the development of techniques that enable a perception of both environments as one. The goal is to transfer and adopt the 3D scene sonification and spatial interaction techniques developed in the last chapter for an application within 3D augmented audio reality scenarios. Therefore, the chapter first reviews the fundamentals of mixed reality applications and defines the principles and requirements of 3D augmented audio reality. In a second part, the chapter develops and implements spatial and location-aware interaction techniques. These techniques are integrated within a self-designed low-cost augmented audio system that is based on an extension of the framework developed in the last chapter.

6.1 AUGMENTED REALITY

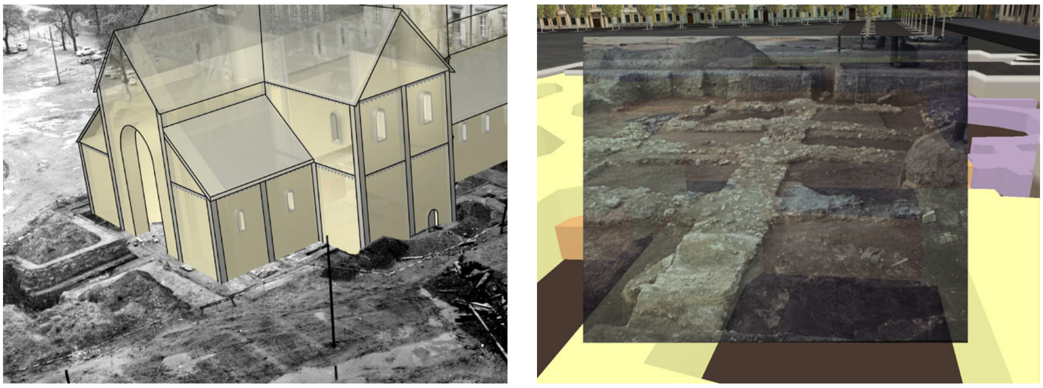
Section 5.1 introduced the term 3D virtual auditory environment and discussed it within the context of virtual reality and the virtuality continuum. This continuum describes the varying degree of virtuality and reality within VR applications, and arranges a real environment on the extreme left and a virtual environment on the right hand side (Milgram et al., 1994), refer to Figure 21 in Section 5.1. The area in between is classified as Mixed Reality (MR) and consists of elements of both environments, real and virtual (Milgram et al., 1994; Azuma, 1997). While the discussion in the last chapter solely focussed on entire virtual auditory environments, this section concentrates on mixed reality and its applications. One of the first definitions of augmented reality – already in respect to the development of augmented *audio* – was provided by Cohen et al.:

> *"Augmented reality is used to describe hybrid presentations that overlay computer-generated imagery on top of real scenes. Augmented audio reality extends this notion to include sonic effects, overlaying computer-generated sounds on top of more directly acquired audio signals. (Cohen et al., 1993)*

The goal of this section is to identify the position of augmented audio reality within the virtuality continuum, and to connect it to the abstract definitions of an enhanced scene $\mathcal{E}$ that were developed in the last chapter.

6.1.1 *Mixed Reality Systems*

The area of mixed reality ranges from Augmented Reality (AR) towards Augmented Virtuality (AV), and is differentiated by the level of virtuality present (Milgram et al., 1994), see also Figure 21. Augmented reality describes a system that enhances and augments a real world environment with additional artificial information (Caudell and Mizell, 1992; Feiner et al., 1993). Augmented virtuality describes the opposite, and defines a virtual environment that is augmented by real world data. The technological and semantical



(a) Augmented Reality. (b) Augmented Virtuality.

Figure 38: Image-based Mixed Reality (Freudenberg et al., 2001a; Röber, 2001).

characteristics are in both systems very similar to pure virtual environments with one major difference. As virtual environments not necessarily resemble a real world place, the (physical) attributes that describe this environment can be arbitrary. This is not true for augmented reality, which is bound to the physical laws that are present at the real world location (Milgram et al., 1995). Accordingly, the term augmented reality has been generalized and redefined by Azuma:

> *"Augmented Reality (AR) is a variation of Virtual Environments (VE). VE technologies completely immerse a user inside a synthetic environment. While immersed, the user cannot see the real world around him. In contrast, AR allows the user to see the real world, with virtual objects superimposed upon or composited within the real world. Therefore, AR supplements reality, rather than completely replacing it." (Azuma, 1997)*

The major goal of mixed reality systems is to present the real and virtual elements in a way that they are perceived as one. Azuma defines three primary attributes for the classification of augmented reality systems (Azuma, 1997):

- Combination of virtual and real elements,
- Interaction in realtime,
- Real and virtual objects are arranged in 3D space.

Examples for augmented Reality/Virtuality.

Two examples for an image-based mixed reality can be seen in Figure 38. In the first example, a virtual reconstruction of the *Magdeburger Kaiserpfalz* is composed over an original photograph of the excavation site (Figure 38a), while the example in Figure 38b shows a photograph of the archeological remains blended over a virtual reconstruction of the excavation site (Freudenberg et al., 2001a; Röber, 2001). According to the definition above, Figure 38a describes an augmented reality scenario, while Figure 38b shows an application for augmented virtuality.

The formal model that was developed in Section 5.1 for describing 3D virtual auditory environments can be easily extended to include augmented audio reality environments as well. The main difference is an additional mapping of the structural information E_S onto a real world location. A bijective mapping can be described as a homomorphism that contains a complete one-to-one mapping of the virtual environment onto a real

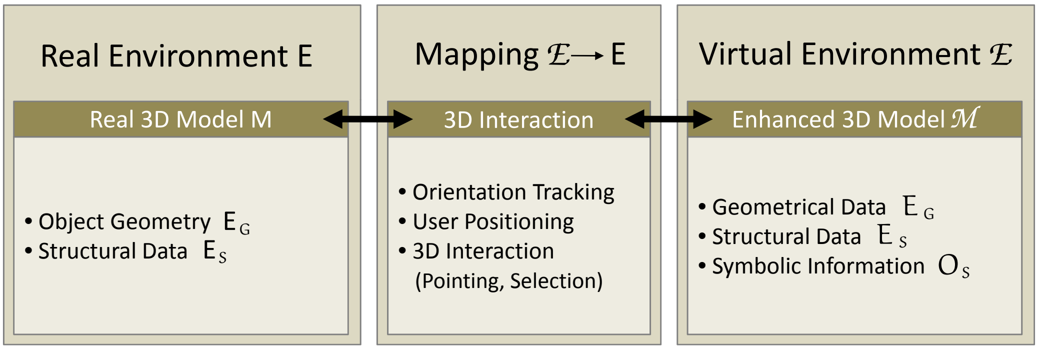


Figure 39: Augmented Audio Reality – Principle.

world place. However, the definition of augmented reality only requires this mapping for at least one object, eg. $\exists \mathcal{M} \in \mathcal{E} : \mathcal{M} \rightarrow M \in \{\texttt{real world}\}$, see also Figure 39. Furthermore, augmented reality enhances and augments not only existing objects, but also adds additional artificial objects and information to a real scene, hence $\mathcal{E} \not\leftrightarrow \{\texttt{real world}\}$. The space and the setting (ie. physics) of the augmented environment, of course, have to match the ones that are present in the real world, as otherwise, the environment will be classified as virtual reality (Milgram et al., 1995).

The definitions of flow, immersion, presence and tele-presence, as were discussed in Section 5.1, are one-to-one applicable to mixed reality systems as well. One major difficulty is the combination of virtual and real objects in a way that they are perceived as existing in one single environment. One important factor is the display quality of the virtual objects, another is the latency of the interaction with the system. To perceive a virtual object as *real*, it has to be presented alike to the real objects in the scene. While computer graphics achieved a high degree of realism at interactive rates in recent years, the latency of mixed reality systems is still an issue (Bowman et al., 2004; Bimber and Raskar, 2005). Latency is thereby a combination of the time required for various tasks. It includes the time for interaction, the time required for the system to respond, as well as the time necessary for the final display of the updated scene. Besides the rendering and display of the output, the interaction devices employed often consume the majority of time (Bowman et al., 2004). The overall latency must be below the update rate of the system, which is for visual applications in most cases around 25fps (ie. latency $\leqslant$ 40ms).

6.1.2 *Augmented Audio Reality*

One of the commercially most successful implementations of an augmented audio reality (AAR) system are *Audio Guides* such as the Sennheiser *guidePORT* (Sennheiser, 2008). These portable devices are employed in museums throughout the world to guide visitors through an exhibition and to display specific information for certain exhibits. Newer systems also feature an automatic user positioning, in which the system automatically detects the user's location by using encoded magnetic fields that are associated with certain exhibits (Sennheiser, 2008). However, as the interaction and user-localization is point based only, these audio guides do not classify as AR system in the strictest sense of its definition (Azuma, 1997).

The fundamentals of augmented audio reality were developed at the same time and together with *visual* AR technologies (Feiner et al., 1993; Cohen et al., 1993; Cohen, 1994). Interesting to note is that augmented audio reality never received very much attention.

In most cases, it is only developed to complement visual AR systems, without a clear focus of the possibilities of an *audio-only* augmented reality. Although an AAR system requires significantly less resources compared to a visual system, difficulties still apply with an efficient and accurate display of artificial sound sources that are superimposed over a real world acoustics. More recently, Härmä et al. describe an augmented audio environment as:

> *"The concept of augmented reality audio characterizes techniques where a real sound environment is extended with virtual auditory environments and communication scenarios. An augmented audio environment is produced by superimposing a virtual sound environment onto the pseudoacoustic environment."* (Härmä et al., 2003)

The acoustic display of the real environment is here integrated into the system using an in-ear microphone/speaker combination. As this especially alters the spatial perception of sound, Härmä et al. describe it accordingly as *pseudoacoustic environment* (Härmä et al., 2003). A more intuitive presentation can be achieved using so called *nearphones*, speakers that are mounted in close vicinity to the ear but without blocking the pinna and the ear canal (Cohen, 1994; Kanno et al., 2006). An interesting alternative are so called bone-conducting headphones, which transmit sound via skin and bone, refer to Figure 45a (Vonia Corporation, 2008). Both systems allow an unaltered perception of the real worlds acoustics, and at the same time a presentation of additional, artificial sound sources.

Auralization of Figure 40.

A small example visualizing an augmented audio reality scene can be seen in Figure 40. It shows the familiar living room environment, this time displayed in a photo-realistic way, with a virtual avatar in its center. In this example, the setting is part of an augmented audio adventure game scenario, in which the main character has to find the exit through the front door. Similar to the examples discussed previously in Chapter 5, the door is initially locked, and the task is to find the key and a certain document to unlock and open the door. The objects which are depicted in red are common 3D sound sources, while the objects in blue are assigned auditory textures, refer to the discussions in Section 5.3.2. The clock on the wall is used as an orientation beacon for the virtual auditory environment. The door's auditory texture conveys in this example the state of being locked, as long as a missing (virtual) document and the key are not found. To achieve this task, several auditory clues are provided as an aid, such as a virtual character (ghost), or additional auditory beacons.

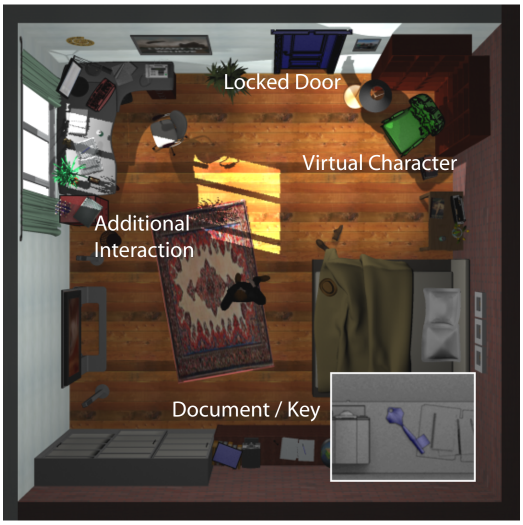


Figure 40: Augmented Audio Example.

An experience of this virtual auditory environment, placed within a real location, can be, depending on its content and implementation, very stimulant and immersive. Although the feeling of presence is directly related to the sophistication of the technologies used, an important factor is the absence of (visual) information, which is now substituted by the user's own imagination. Similar to the reading of books and the listening to radio plays, it can be assumed that a user's immersion and the feeling of presence within an augmented *audio* reality is stronger than in any other VR/AR system. Although this statement is difficult to prove, Chapter 9 discusses and evaluates several applications that are concerned with this statement.

Using these discussions and the descriptions of Azuma and Cohen et al. (Azuma, 1997; Cohen et al., 1993), augmented audio reality is defined for the use in this research as:

Definition *3D Augmented Audio Reality* describes a system and techniques that allow an enrichment of an existing real world environment with additional auditory data and information. The system is thereby centered around an auditory design, which describes a strong focus on the benefits of an auditory display and an audio-only application. The superimposed 3D virtual auditory environment thereby complies with the real world place in terms of (physical) characteristics and topology. Spatial interactions are performed similar as to 3D virtual auditory environments, with an added exploration through a real world user positioning and orientation tracking.

The requirements for a 3D augmented audio reality environment are:

- Based on a (non-realistic) 3D auditory display design.
- Employs efficient techniques for user localization and orientation tracking.
- Uses a realtime 3D spatial interaction design (positioning, pointing).
- Is developed around an auditory-centered design.

Requirements

The requirements for the development of an AAR system are in many aspects very similar to the demands of a 3D virtual auditory environment. The most prominent difference is additional user-tracking and positioning equipment, which is required for a localization/orientation of the listener within the real/virtual environment. According to the above provided definition and the discussions thus far, the requirements for the development of an augmented audio reality system are:

- A 3D virtual auditory environment framework with suitable:
  - Task-related sonification techniques,
  - Task-related spatial interaction techniques, as well as
  - A non-realistic 3D auditory display.
- A proximaural sound presentation system.
- User positioning and orientation techniques (tracking equipment).
- 3D spatial interaction devices.
- Realtime sound spatialization and acoustic simulation techniques.
- All integrated within a lightweight and transportable system (wearable computer).

Due to a less sensitive auditory perception, compared to vision, the requirements for the tracking accuracy are not as high as for visual applications. It can be assumed that, depending on the task and application, a tracking accuracy of +/- 2m for the listener's position is sufficient. This, however, is only true if the setting and the objects for interaction are large enough. In small environments, such as an office, a different positioning approach is required, and can be realized using Bluetooth and other proximity-aware technology. Using an intelligent design for the augmented environment, the required accuracy can further be *decreased* to allow the employment of free available, everyday technology. The accuracy for measuring the user's orientation (head-tracking), however, has to be as high as for 3D virtual auditory environments.

The remainder of this chapter will analyze and discuss these requirements in more detail, and extend the audio framework of the last chapter towards an employment for augmented audio reality applications. Section 6.2 concentrates on the required location-aware interaction techniques and discusses both, tracking hardware as well as the methods required for a 3D spatial interaction design.

Applications and Prospects

Several implementations of augmented audio reality already exist and were beneficially applied in many areas. These range from entertainment & edutainment applications, over the guiding of visually impaired, to the support of daily tasks using ambient intelligence and ubiquitous computing (Cohen, 1994; Rozier et al., 2000).

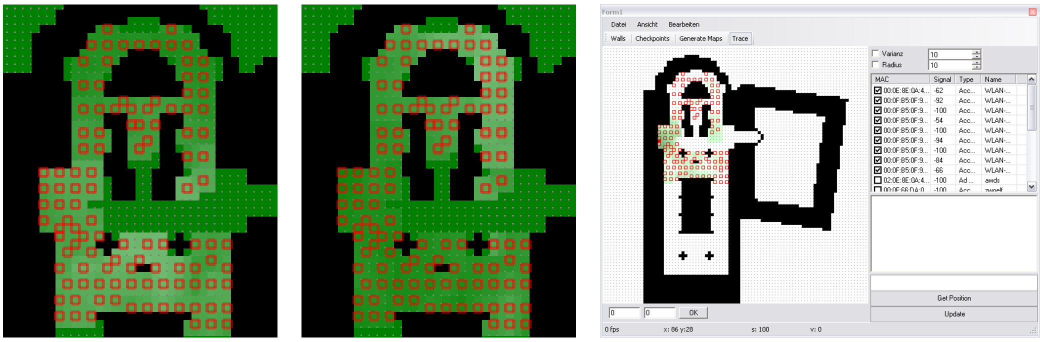
Similar to virtual reality, the possibilities of augmented (audio) reality systems inspire and are explored by artists and scientists together: Augmented audio is most often employed for an exploration of exhibits in a museum (Sennheiser, 2008), for an acoustic examination and exploration of history in the *Berlin Wall* project (Mariette, 2006), as well as for playing computer games (Cohen et al., 2004). Mariette has designed several exhibitions that focus exclusively on the possibilities of augmented audio reality and the perception of a personal, location-sensitive 3D soundscape (Mariette, 2006, 2007a). Due to the less rigid hardware requirements, some implementations of augmented audio on portable devices, such as mobile phones, exist (Ekman, 2007).

The technology of augmented audio reality is in general very well suited to aid guiding and navigation related tasks. Several prototypes – most of them developed for aiding the visually impaired – have been designed to simplify and enhance a navigation and orientation in complex environments, such as cities and large office complexes (Cohen, 1994; Härmä et al., 2003; Walker and Lindsay, 2005). Interesting would be also an application of augmented audio reality to support the performance of daily tasks in the form of audio-centered ambient intelligence and ubiquitous computing. As current developments continue to increase the visual complexity of our environment, augmented audio might be a more suitable technique for a non-intrusive display of data and information.

6.2 LOCATION-AWARE INTERACTION CONCEPTS

A key aspect in augmented reality applications is the possibility to interact with the system within a real-world setting. For this purpose, efficient and accurate user-tracking devices and techniques are required to determine the user's position and orientation. Using these two techniques, the user is able to navigate and orientate oneself within an augmented environment in the search for places of interest for further exploration. At these locations, spatial interaction techniques are required to interact with virtual and/or real objects, to gather more information and knowledge (Bowman et al., 2004; Bimber and Raskar, 2005). Although common interaction techniques using a computer keyboard and mouse would be sufficient in many cases, a spatial interaction design has the advantage of emulating interactions from the real world, and thereby directly improves the perception of both environments as one.

The following two sections are therefore dedicated to the development of efficient and cost-effective user tracking and positioning techniques, as well as to the design and exploration of spatial interaction metaphors. The goal of this section is not only to device



(a) Radiomap AP2. (b) Radiomap AP5. (c) Positioning Application showing the User's Location Probability (green).

Figure 41: WiFi-based User Positioning.

an augmented audio reality system and techniques in accordance to 3D virtual auditory environments, but to design a framework that is efficient, reliable and uses commonly available low-cost components.

6.2.1 *User Positioning and Tracking*

The problem of *registration* is one of the central aspects in augmented reality, in particular for visual AR systems. Dedicated user tracking equipment can thereby also be used for a detection of gestures and movements to permit an intuitive interaction with the virtual environment (Rose et al., 1995; Azuma, 1997). Rolland et al. describe the importance of user tracking and positioning techniques for AR systems as follows:

> *"Tracking for virtual environments is necessary to record the position and the orientation of real objects in physical space and to allow spatial consistency between real and virtual objects." (Rolland et al., 2001)*

Different systems for measuring the user's position and orientation are available, and include inertia-based sensors, time-of-flight systems, mechanical devices, as well as spatial scans and phase-difference sensors (Rolland et al., 2001), see also Figure 42 for some examples. As the focus of this research is the development of a cost-effective system, several of theses technologies already disqualify. The global positioning system (GPS), which is classified as a time-of-flight system, is employed in a few outdoor-based augmented reality scenarios (Cohen et al., 2004). However, GPS systems can not be used indoors and also exhibit a positioning accuracy of only 8–10m (Rolland et al., 2001). Other cost-effective user positioning techniques are based on Bluetooth, or the WiFi radio network (Youssef et al., 2003). An advantage of using Bluetooth is that it can be employed easily and is also very stable and reliable. The disadvantage of this technology is that it only permits, similar to Sennheiser's guidePORT system, a point-based user localization, which does not allow a complete and continues tracking of the user's position (Härmä et al., 2003; Sennheiser, 2008).

A video explaining Figure 41.

A positioning technique that enjoys a continues increase in popularity and research is the so called *WiFi Radio-based Positioning* (Youssef et al., 2003; Youssef and Agrawala, 2005; Ivanov and Schemmer, 2007). This technique uses publicly available WiFi access points to determine a user's location. In a pre-processing step, a radiomap is measured and generated which displays the radio signal strength for each access point and location

(a) InertiaCube3[1] (InterSense, 2004).

(b) OceanServer OS5000-US Compass[2] (OseanServer Technology Inc., 2007).

(c) Gyration UltraGT[3] (Gyration, 2004).

Figure 42: 3D Interaction and User-tracking Devices.

in the map. During the later user tracking, this information is compared to the current signal strength, in which the system generates a probability map to determine the user's current location (Youssef et al., 2003; Youssef and Agrawala, 2005; Otto and Domke, 2007). An example can be seen in Figure 41. Here Figure 41a and Figure 41b display radiomaps for two access points that were used in an augmented audio reality experiment that took place in the Magdeburg Cathedral, refer to Section 9.5 for a more detailed discussion. In this experiment, a total number of 9 access points were used, see also Figure 43. Figure 41c shows a screenshot of the application during the tracking procedure and the determined user's location, which is highlighted by a red square. An advantage for using this technique is the growing number of publicly available WiFi access points, as well as a large variety of WiFi equipped mobile hardware, such as mobile phones, laptops and portable game consoles (PSP).

As the localization is based on signal strength alone, it does not interfere with any security issues and the positioning also works properly with encrypted access points. The accuracy of the positioning varies from 2–6m and is dependent on the environmental geometry, the total number of access points used, as well as on the quality of the pre-measured radiomap (Röber et al., 2006a; Otto and Domke, 2007). This accuracy is sufficient for augmented *audio* reality applications if the setting is distributed over a larger area. The position, however, has to be interpolated and additional measurements have to be employed to avoid ambiguities. As this radiomap is one of the key aspects, an efficient and accurate simulation of radio signal propagation is imperative. Similar to other areas, computer graphics hardware can here be employed to improve the simulation's quality and especially its efficiency (Rick and Mathar, 2007). A close examination of Figure 41a and Figure 41b exhibit a shadowing effect of the radio signal that occurs behind walls and large obstacles.

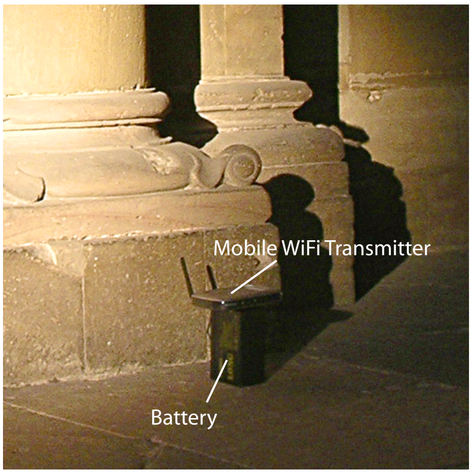


Figure 43: Mobile WiFi Access Point.

1 http://www.isense.com/
2 http://www.ocean-server.com/
3 http://www.gyration.com/

As of this effect, a WiFi-based user tracking is of higher accuracy indoors than outdoors (Röber et al., 2006a; Skyhook Wireless, 2008). For a free field experiment, the access points have to be arranged in a way to create an artificial signal shadowing, ie. that they do not evenly cover the entire area. Newer measurements in Berlin, with a very high number of simultaneously *visible* access points ($\geqslant 25$), suggest a constant positioning accuracy between 1–2m (Otto and Kurth, 2008).

Just as the user positioning, the measurement of the user's head orientation is of high importance as well and has to be performed with a high accuracy and a low latency. Several devices to accomplish this task are depicted in Figure 42 (Rolland et al., 2001). The example in Figure 42a is a rather expensive device, which is based on high accuracy inertia sensors, while the device in Figure 42b shows a digital compass and the example in Figure 42c is a 3D mouse based on a gyroscope (InterSense, 2004; OseanServer Technology Inc., 2007; Gyration, 2004). All devices provide a very high resolution of up to 0.1° and have a latency that ranges between 20–50ms.

For a low-cost implementation of an augmented audio reality system, the digital compass and the gyro-based 3D mouse have the highest applicability. Both devices can be employed in a wearable computer setup and do not require additional hardware or sensor equipment. Also, the WiFi-based user positioning appears to be well suited for augmented audio reality, as this technology can be applied to indoor and outdoor settings. However, care has to be taken in the arrangement of the access points and the measurement of the radiomap.

6.2.2 *Spatial Interaction*

The last section discussed several techniques and hardware devices, which are applicable to perform 3D spatial interactions within virtual/augmented environments. A direct interaction with virtual objects, eg. to acquire information and/or to select/activate objects, requires spatial interaction techniques similar to those discussed in Section 5.4. The focus of this section lies on the development of techniques that efficiently combine both worlds, ie. to define an interaction on virtual objects that are based in a real environment.

An interaction with augmented environments is very similar to the interaction design of entirely virtual spaces. Basic tasks to be performed are selection, manipulation and/or activation of objects, as well as a recall of information. The main difference, compared to an interaction with pure virtual environments, is an interaction in real world space. This requires a precise and congruent mapping of both – virtual and real – environments in order to be perceived as one. The WiFi-based user tracking technique that was discussed in the last section provides here an average positioning accuracy of about 3–4 meters, although experiments with high-quality radiomaps suggest a possible accuracy of around and less than 1 meter (Youssef and Agrawala, 2005; Otto and Kurth, 2008). A tracking accuracy of 4m is sufficient if the augmented audio environment only contains large objects that are adequately separated

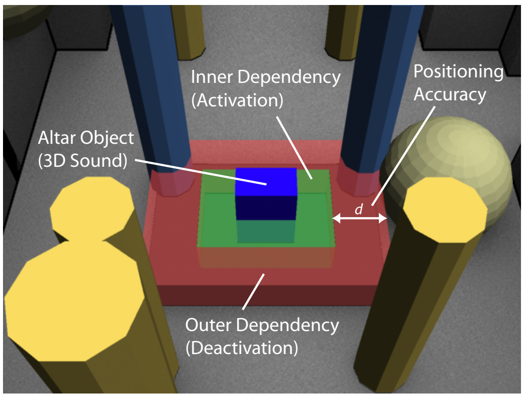


Figure 44: Design for Position Dependency.

from each other, refer to Figure 44 and Section 9.5 for a discussion of examples. The position dependency, as it was introduced in Section 5.5, describes a user controlled interaction with the environment based on proximity. To compensate a less accurate positioning performance additional measurements are required, as is shown in Figure 44. These consist of two bounding boxes of different size, which act as switch and either activate or deactivate a certain object. The inner bounding box is thereby used to activate a position dependency, while the outer bounding box is used to switch it off. The differences in size between the bounding boxes d must be large enough to compensate for the positioning error, ie. for the WiFi based positioning around 2–3 meters. This approach, however, is only applicable to larger settings with sufficiently separated objects. An alternative implementation that circumvents these limitations can be realized through additional proximity sensors, such as Bluetooth and RFID technology. Here an object is activated as long as the user's distance is below a certain pre-defined threshold. The implementation for the other environmental dependencies, eg. time- and object-dependencies, can be directly taken from 3D virtual auditory environments without any modification.

All spatial interactions that were devised in Section 5.4 are directly applicable also for an interaction with augmented auditory environments. The employed 3D interaction devices, however, have to be both, portable and wireless. Although a regular wireless gamepad can be used to model the majority of interactions, an application of a gyro-based 3D mouse is here much more interesting, refer to Figure 42c. Such a gyro mouse is employed in the implementation of this framework as a 3D interaction device. With two buttons and an additional scrollwheel, this device can be well employed in a large variety of 3D spatial interaction scenarios, such as for 3D object selection, pointing and picking, but also to interact with 3D auditory menu systems.

After these discussions, the following section reviews the previously devised audio framework and extends it towards an application for augmented audio reality scenarios. The here discussed techniques for user tracking and positioning, as well as for 3D spatial interaction are thereby integrated into the system.

6.3 SYSTEM DESIGN

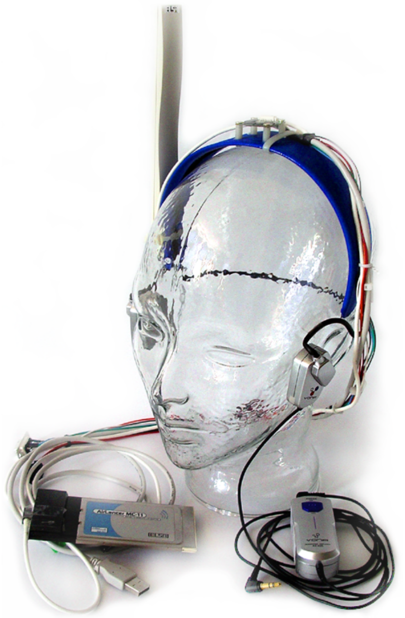
The audio framework of Section 5.5 can easily be extended towards an application for augmented audio reality. The augmented audio reality system designed is meant to be used in a variety of applications, ranging from pure enter- and edutainment, to the development of applications to aid the visually impaired. Several experiments and case studies have been performed and are presented and discussed in more detail in Chapter 9.

The development of the system is divided into hardware-related and framework design issues. The first section describes the selection of specific hardware for 3D user interaction, orientation tracking and positioning, while the second part discusses the modifications of the framework regarding an implementation of the 3D spatial interaction techniques.

6.3.1 *Hardware Requirements*

A major goal for the design was the development of an affordable system based on efficient, but low-cost components. The hardware requirements therefore are:

- A portable and lightweight computer platform (Laptop or PocketPC), with
    - Realtime sound spatialization and acoustic simulation hardware.



(a) User Positioning and Tracking, Bonephones.

(b) 3D Pointing and Selection.

(c) Play and Interaction.

Figure 45: Augmented Audio Reality – System Hardware.

- A proximaural headphone system,
- User-orientation tracking hardware (head-tracking),
- A user-localization/positioning system, as well as
- A wireless 3D spatial, and a standard (gamepad) interaction device.

An overview of the selected hardware can be seen in Figure 45. It shows in Figure 45a a digital compass for the head-tracking, a special antenna for the WiFi-based user-positioning, as well as bone-conducting headphones as a proximaural display. Figure 45b displays the 3D interaction device, a common gyro-based 3D mouse, while Figure 45c shows a regular wireless gamepad for the input of standard interactions. Additionally, several mobile WiFi access points were employed for the experiments and for a test of the system, refer to Section 9.5.2. These access points were dispersed manually within the environment to ensure a good coverage, but to also introduce signal shadowing artifacts of pillars and walls, refer also to Figure 41.

The selection of this hardware was made for several reasons. The components exhibit all of the required qualities and are additionally highly affordable. The bone-conducting headphones are available for € 100 and can be well employed as an acoustic display in auditory environments (Vonia Corporation, 2008). The results of a short analysis and comparison with a regular HiFi headphone system are available in Section 9.5.1. The digital compass employed in this system is available for € 200, possess several serial interfaces and achieves an update rate of 40Hz ($\approx$ 25ms) with an accuracy of 0.1° (OseanServer Technology Inc., 2007). The WiFi-based user tracking is performed using an external, but regular WiFi computer card that can be equipped with an additional antenna and is available for € 60. The 3D pointing and interaction device that was employed is a gyro-based computer mouse, which is available for around € 50.

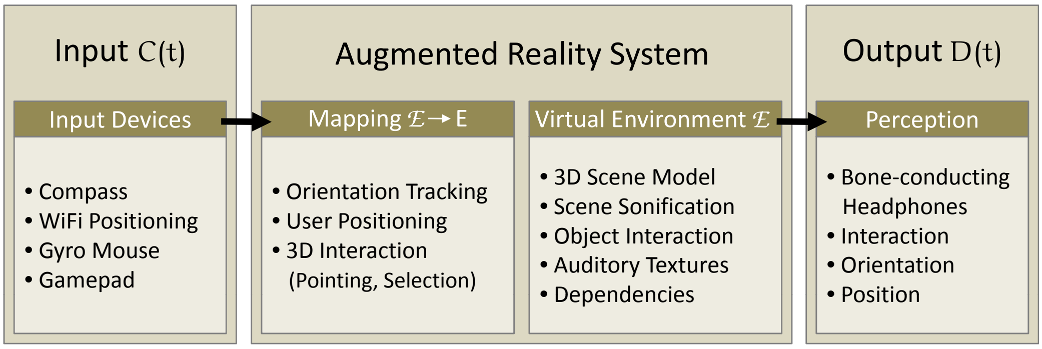


Figure 46: Augmented Audio Reality Framework.

6.3.2 *Framework Extension*

The system itself is an extension and implemented on top of the audio framework that was devised earlier in Section 5.5. It is implemented using C ++ and oriented along the requirements of mobile hardware, ie. laptops and PocketPCs. Although the framework's performance on less efficient hardware, such as a PocketPC running WindowsMobile, has not yet been evaluated, it should perform with an adequate efficiency. As in the previous implementation, the system employs OpenSG as 3D scenegraph system, OpenAL for sound rendering 3D spatialization, as well as OpenGL for an additional 3D visualization of the scene. A freely programmable DSP-based sound API for sound rendering and sound spatialization has already been realized and is available for an application on mobile systems (Stockmann, 2007). Interesting to note is that this API is more efficient than any other sound API available and also allows an exchange of the HRTF filters employed for a personalization of the 3D sound spatialization.

The majority of aspects that were discussed previously in Section 5.5 are directly applicable to an implementation of augmented audio reality as well. An overview of the augmented audio reality system developed can be seen in Figure 46. Based on the original audio framework, an additional mapping of the 3D virtual auditory environment $\mathcal{E}$ onto a real location is performed. This mapping is performed through several input devices and techniques that allow an efficient user tracking and positioning. Additions to the audio framework therefore include a WiFi-based positioning system, as it is depicted in Figure 41, as well as a support for the new head-tracking and 3D interaction devices, refer to Figure 45. Figure 41 shows a screenshot of the developed WiFi-based user positioning system, which allows to determine the user's position within a real environment with an accuracy of about +/- 2m. This system is implemented using C ++ as well, and integrated into the main system as external library (DLL). The digital compass, which is employed as 3D head-tracking device, is connected to the system using a regular serial interface. With an update rate of 40Hz, the compass communicates with the system and transmits three angels: heading, pitch and roll, which are used to determine the user's head orientation. The main interaction with this system is performed using positioning, ie. position dependencies, as well as through a control via a regular gamepad and a 3D pointing and selection device. For the 3D pointing and object selection, a Gyro mouse is employed, which is controlled using a standard mouse device that converts the 2D mouse coordinates into angels for a selection and picking of virtual 3D objects. Feedback is perceived via bone-conducting headphones (auditory output), refer to Figure 45, as well as through the user's position and orientation within the real environment. Section 9.5.1 examines the bone-conducting headphones in more detail and compares the perception of 3D virtual sound sources and virtual room acoustics using bonephones and regular HiFi headphones.

6.4 SUMMARY

After the discussions of 3D virtual and augmented auditory environments, the areas of application and the design of actual examples and prototypes moves into the focus of this research. Figure 47 in the following Chapter 7 provides an overview of the entire system and shows the integration of an authoring component within the current audio framework. Additionally, this chapter discusses issues regarding the authoring and design of general applications that are based on 3D virtual/augmented auditory environments. The chapter develops rules and guidelines for various authoring tasks and devises a 3D authoring environment that is used to implement several examples, refer also to Section 9.5.