7 AUTHORING AND DESIGN

I paint things as I think of them, not as I see them. — Pablo Picasso

IN addition to the representation and display of 3D virtual and augmented auditory environments, the design, as well as the process of authoring these environments is of high importance. This not only includes the design, selection and assignment of sounds and sound sources to certain positions, but also the definition of interactions, dependencies, auditory textures, as well as the setup of an underlying *context* – the symbolic information E_S. The goal of this chapter is therefore to identify and describe the authoring process for 3D virtual auditory environments, as well as to derive general authoring paradigms and guidelines that are suitable for the design of a large variety of applications and tasks.

The first section of this chapter reviews common design principles for 3D auditory environments. Additionally, the section examines existing approaches for the authoring and design of sound and acoustics as they are employed in entertainment computing, such as for the design of audio/visual computer games. The goal of this discussion is the development of an authoring pipeline that supports the design of 3D virtual/augmented auditory environments. The discussion continues in the following section by providing authoring guidelines and design aspects. The process of authoring is thereby divided into smaller tasks, which are analyzed individually and for which several principles and guidelines are presented. The chapter concludes with the design of a prototypic authoring environment, which is applied to the authoring of an example scene to show some of the developed techniques in practice.

7.1 AUDITORY AUTHORING

Authoring describes the process of creation and the adding of content for a specific medium. Authoring environments are able to aid this process and are employed in many areas to *create* the actual application. A related area to the authoring of 3D virtual/augmented auditory environments exists in the form of designing audio/visual computer games, and in particular, the design of interactive 3D adventures. Despite many rumors, the adventure genre is still very active and kept alive by a large community, who also creates and develops own authoring systems for the design of 2D and 3D adventure games. Outstanding examples are the *3D Adventure Studio*, the *Adventure Game Studio* and the *Adventure Maker* (van der Honing, 2003; Jones, 2008; Giovanni, 2008).

A commonality in all these systems is their room-centered design approach, which allows the authoring and an easy connection between different rooms and environments. The design of a 3D room/environment includes a selection of background graphics, masks and sprites for walkable areas and walk-behinds, hotspots and objects for interaction, as well as a selection of music and sound effects. Through the design of several of such rooms and an integration with story-related users tasks an adventure game is created. Secondary tasks of the authoring include the design of a user interface, the development of an asset management system, as well as the authoring of content, narration and user interaction techniques. The user interface thereby combines all other applications and later manages and controls the player's progress and play. Interesting to note is that the majority of authoring environments provides an integrated runtime component,

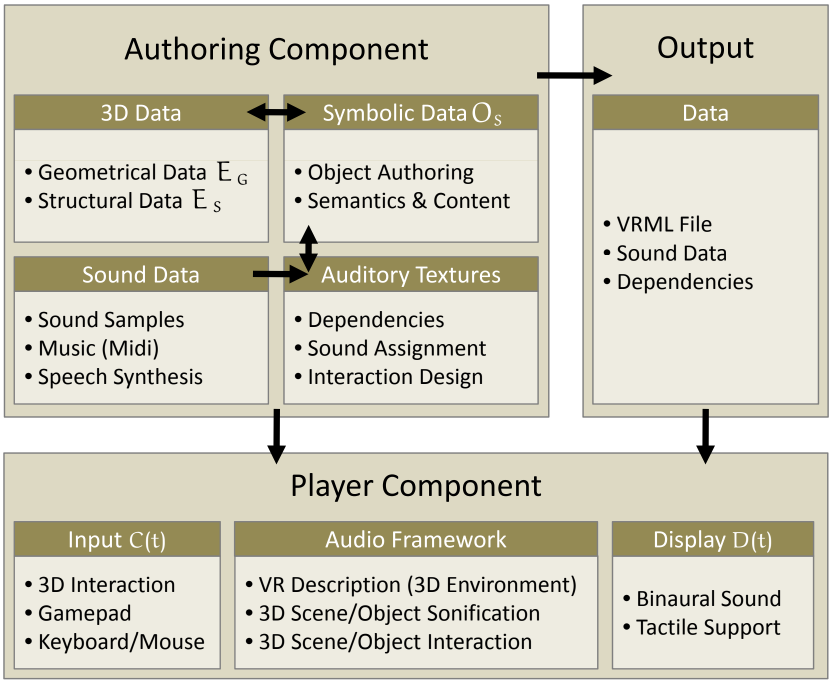


Figure 47: Audio Framework – Overview.

which allows a comfortable preview and evaluation of the game directly from within the authoring application.

The authoring process for 3D virtual auditory environments is not much different than the design of a 3D audio/visual adventure game. However, certain aspects, such as the auditory presentation and perception are quite different and have to be considered during the entire authoring process. With a consideration of an authoring for 3D augmented audio applications as well, also the hardware and sensors employed move into focus. One requirement of the authoring system is therefor to define an interface between the runtime system and the 3D interaction hardware used, see also Figure 47. This is especially important for the design of augmented audio reality applications, which rely heavily on external hardware for user interaction and positioning. As the hardware of AR systems is not yet standardized, a related authoring environment is often only applicable to this specific system (Paelke and Reimann, 2006). The required components for the authoring of 3D virtual/augmented reality applications are described by Paelke and Reimann as:

- Surveying and capturing of an environment for augmentation,
- Modeling of content for augmentation,
- Definition of dynamics (eg. for user interaction), as well as
- Spatial registration of both (real and virtual) environments.

This list, although its authored applications focus on a *visual* perception, can also be employed for the design of an authoring system specific to the creation of a 3D *auditory* environment. Such an audio-centered authoring and design is a non-trivial task and includes much more than just the assignment of sound sources to certain positions in 3D space. The process of authoring itself is founded and based on several design guidelines that were derived during this research, and whose discussion is the focus of the succeeding section.

An overview of the entire audio framework, and the authoring process in general, can be seen in Figure 47. It shows that the framework is basically divided into two parts, an authoring component and a runtime system. The authoring starts with the creation of a 3D model that defines the augmented environment and resembles a real-world place. The 3D model is described by geometric data E_G in the form of polygons and triangles, as well as through structural information E_S that defines the topology and the mapping of the 3D model onto the real-world environment. The second step comprises the design and creation of sounds and sound effects, as well as speech and music samples, which are later added to the virtual environment and assigned to certain objects. Thus far, the authoring was performed using external tools and applications, such as *3DStudioMax* for 3D modeling and *Soundforge* for sound design (Autodesk, 2008; Sony Creative Software, 2008), see also Figure 50. After this *pre-authoring*, the design of the actual 3D virtual auditory environments is performed through the definition and assignment of object-related symbolic information O_S to each object within the scene. The result of this authoring is the transformation of a regular 3D scene E into an acoustically enriched 3D enhanced environment $\mathcal{E}$.

The transformation (authoring) is mainly based on the design of time-, object- and position-dependencies, the creation and assignment of 3D sound sources, as well as includes a selection of task-related sonification and interaction techniques and the definition of auditory textures for the virtual scene objects, refer to Figure 47. After the authoring is complete, the acoustically enhanced environment $\mathcal{E}$ is saved along with all data and scene information within an extended VRML data file. This data can later be loaded into the runtime system, and be used for playback and to explore and experience the authored 3D virtual auditory environment.

Within the context of this research, the authoring and design of 3D auditory environments is defined as follows:

Definition The *Authoring and Design* of a 3D virtual/augmented auditory environment is described as the transformation of a regular 3D scene E into an acoustically enriched 3D enhanced environment $\mathcal{E}$. The authoring is divided into a pre-authoring step, which includes the creation of scene geometry E_G and an authoring of structural information E_S. The actual 3D auditory scene design performs a definition and assignment of symbolic data O_S to the entire scene and to individual objects to design enhanced 3D models $\mathcal{M}$. The authoring pipeline comprises:

- The creation of geometry and a virtual (geometric) 3D scene authoring,
- The design of sound, speech and music samples,
- The definition of dependencies and a selection of sonification and interaction techniques, as well as
- A scene object authoring through the definition and setup of auditory textures.

Besides this technical creation of 3D virtual auditory environments, also the arrangement and authoring of sounds and objects within this environment is crucial and requires a careful design. A well authored application is here always able to intuitively convey enough information to the listener. A badly designed environment might contain empty and silent areas in one place, while it overburdens the listener with too much information and too many sources in other locations.

7.2 AUTHORING GUIDELINES

Several parts of the authoring process are task-related and can vary depending on the content and intent of the final application. An entertainment application, such as an audiogame, has different requirements than an augmented audio application that aims to aid the navigation and orientation of the visually impaired. The similarity of both examples is the underlying authoring pipeline, as well as the design guidelines and authoring techniques used. As discussed in the previous section, the basic authoring is similar for all applications and requires geometrical data, structural information, sound, speech and music samples, as well as content and symbolic information. The differences reside in the authored environments themselves, more specifically in the 3D models, the intent, the content, the selection of sonification and interaction techniques, as well as in the later display and use of the final application.

Chapter 4 described several examples for 2D and 3D auditory displays along details of their implementation and design principles. The majority of these principles are also applicable and can be used for the design and authoring of 3D virtual auditory environments as well. Despite a task-related authoring, some issues are common in the design of all applications. This includes primarily a careful selection and design of sounds and samples to support an expressive and efficient sonification of the environment. But it also includes the principle that not all objects can be audible at the same time, as this would clutter, and render the auditory display inoperable. A procedure has to be employed that controls the playback of sound sources and sonifications consistent to the listener's interaction. Auditory textures can here be used together with distance attenuation techniques to blend out farther away sources, as well as to activate sound sources using proximity and a positional dependency.

To exemplify the process of authoring and to discuss some of the authoring techniques, the familiar 3D environment is employed in the setting of an augmented audio reality game. Figure 48 provides here an overview, as well as shows the required authoring steps to create a small audio-based AR game. In accordance to the previously discussed examples, the environment resembles a part of an augmented adventure game, in which the player's first task is to find an exit to this room. The door, however, is locked, and the player needs to find a key to unlock and open the door. Therefore, both objects, *Door* and *Key*, are assigned auditory textures that describe these objects, as well as implement possible interactions. The door object represents two states, locked or open, while the key emits a descriptive sound that identifies itself as key. Both objects have a position dependency, in which the key's position dependency also activates an object dependency that unlocks the door, refer also to Figure 37. The key object additionally features a time dependency, which is used as an aid for the player to find the key. A further authoring for this scene is required to setup the environment and to provide additional auditory feedbacks, eg. environmental sounds, a narrators voice and music. The clock on the wall

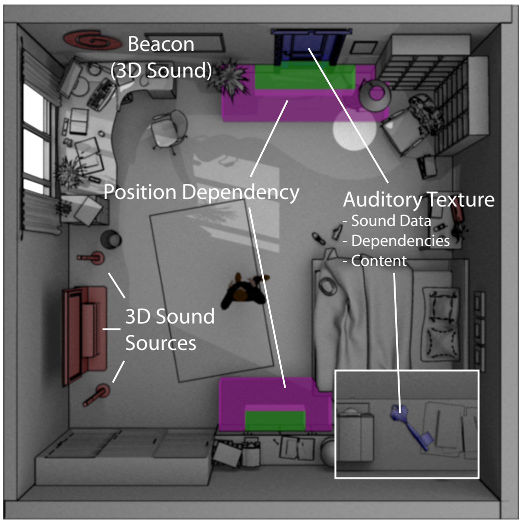


Figure 48: Authoring Example.

can be employed as a an auditory landmark and is used as a *north beacon* to provide additional cues for the user's orientation. These tasks require the specification, setup and positioning of 3D sound sources, as well as the use of dummy objects and additional auditory textures to implement a more detailed narration.

The authoring of the example depicted in Figure 48 is relatively easy and straightforward. The following three sections discuss specific problems within the authoring pipeline, as well as devise guidelines and principles to improve the design process.

7.2.1 *3D Scene Authoring*

3D scene authoring is concerned with the actual physical design and 3D modeling of the virtual environment. This includes the design and creation of individual 3D models and their arrangement within a larger topology. The designed 3D models do not need to be very realistic, as they are mainly used as placeholders for an assignment of auditory textures, as well as for collision detection. Therefore, in most cases simple boxes, spheres and cylinders are sufficient. This allows a fast design and modeling of scene geometry and also to perform changes and adjustments quickly and without many difficulties. Furthermore, low resolution models require less storage space and can also be rendered more efficiently using OpenGL for a possible *visual* scene display. Care must be taken in the positioning of virtual objects for an augmented audio reality application. Depending on the positioning technique employed, objects with a position dependency have to be separated by at least twice the positioning accuracy, refer to Section 6.2. 3D scene authoring also comprises the design and compilation of speech, sound and music samples for their later use and assignment to individual scene objects. Although high quality sounds are required in all cases, some applications may have secondary requirements, such as storage space in mobile applications, or to account for perceptional differences due to the use of a bone-conducting headphone system. For the example depicted in Figure 48, the 3D scene authoring includes an approximate modeling of the room's interior, as well as the compilation of a sound pool for later sound assignments. 3D models can be created using tools such as 3DStudioMAX, from which they can be exported and stored as VRML data file, refer to Figure 50. Objects in this VRML file are used in the next authoring step, in which they are extended by the integration of sound nodes and auditory textures.

7.2.2 *Content and Dependency Authoring*

The second step in the authoring pipeline is the creation of content; the authoring of the very essence of an auditory environment. A large variety of applications can be implemented just through the modeling of various position, time and object-related dependencies. Their application is very versatile and can also be employed in the form of a *state-machine*, in which additional *dummy objects* are used to store boolean variables for the various states. In this setting, and by also utilizing time-dependencies, a state-machine can be employed as story-engine for the construction of a story arc. The boolean variables thereby contain the various story elements, which are activated either using a time-, or user-triggered object dependency. In the above example, a time dependency is added to the key's auditory texture, which emits a hint after a certain amount of time has passed. An object-based dependency is added to the key object as well, which unlocks the door after the key has been activated. Both, the key as well as the door object contain a position dependency as it is shown in Figure 48. This position dependency activates both objects and acoustically describes their current state. If the key object has been previously

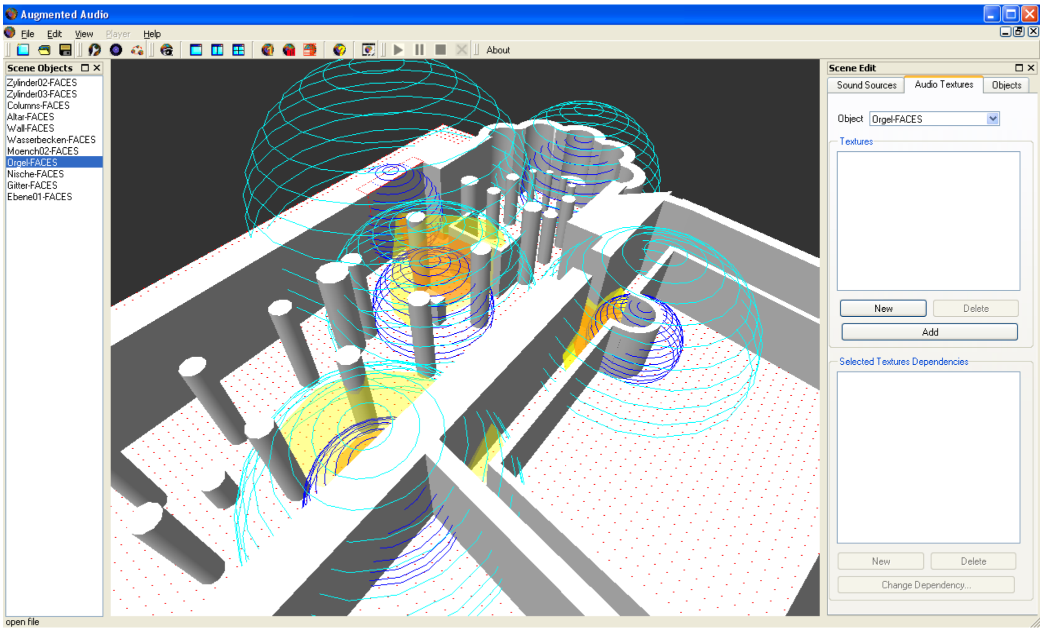


Figure 49: Auditory Authoring Environment.

activated, the door opens, and the player can proceed to the next level. Additional content and story data, which is not shown in this example, can be added to provide a denser storyline and a more realistic atmosphere. This can also be authored simply through the creation of additional 3D sound sources and the assignment of auditory textures.

7.2.3 *Sonification Design*

After the story authoring is complete, the existing auditory textures are extended to implement user interactions and additional object sonifications. The wall clock in Figure 48 is employed as a scene-embedded north beacon, and assigned a sufficiently loud ticking sound. The ticking sound is spatialized with an omnidirectional emittance pattern (point sound source) to be used as an orientational aid. The remaining auditory representations for the key and door object can now be added as well, and might be additionally enhanced by a narrative speech-based description. Finally, all remaining objects are authored and auditory textures are used to define the interaction with these objects and the sonifications used. Music and the narrator's voice are assigned to dummy objects, for which object and time dependencies are used to control their playback.

7.3 DESIGNING AN AUTHORING ENVIRONMENT

An authoring environment that supports the previous discussions and authoring concepts can be developed in accordance with the existing audio framework that was devised in Section 5.5 and extended in Section 6.3. An overview of the entire framework was provided at the beginning of this chapter in Figure 47, which highlights its two major components: an authoring and a runtime system. The runtime system is thereby fully integrated into the authoring component and allows a direct preview of the authored scenes. The authored applications are saved as an extended VRML data file, and are later loaded together with the sound and music samples into the runtime system for playback.

The runtime system can thereby be specific and focus on a certain type of application only, such as 3D audiogames or an augmented audio reality system, but the authoring itself shall be performed in a single, task-overlapping authoring environment, see also Figure 49.

The actual authoring and design starts with an existing 3D model of the environment and a large sound pool of representative sound, speech and music samples. The requirements for developing an authoring system that supports all of the previous discussions are:

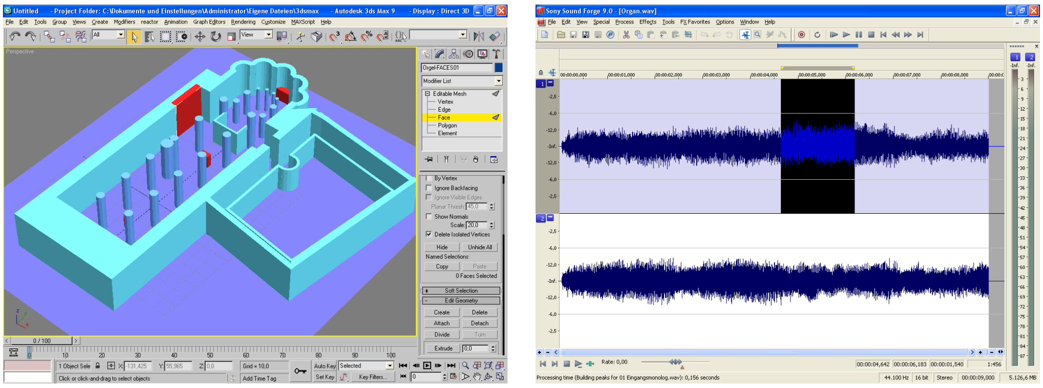
- General Application
  - Load/save VRML data files
  - 3D scene visualization (OpenGL)
  - 3D Sound and object visualization
  - Dummy object creation and positioning
  - General environmental acoustics authoring
  - Listener authoring (start position, ...)
  - Menu and mouse based interaction
  - Scene authoring preview

- 3D Sound Source Authoring
  - Select, create and delete sources
  - Sound assignment
  - 3D Alignment and positioning
  - Parameter adjustment (gain, filtering, spatialization, ...)
  - Distance attenuation

- Auditory Texture Authoring
  - Select, create and delete auditory textures
  - Object assignment
  - Time-, object- and position-dependency authoring
  - Sound assignment and individual parameter adjustment
  - Interaction selection (input dependency)

The list differentiates between general authoring requirements, and specific needs for the creation and design of 3D sound sources and auditory textures. A graphical display of the 3D environment is required, as well as a visualization of the individual authoring tasks. This allows a much easier scene design and parameter adjustment. Examples are shown in Figure 49, but are also discussed in more detail in the following sections.

This authoring approach is only suitable for the design of 3D auditory environments, and in this form not applicable for the design of 2D auditory displays and non-topological environments. Section 9.6 later introduces such a *non-spatial environment* in the form of Interactive Audiobooks, and thereby also describes a system that focusses on the authoring of these applications.



(a) 3D Scene Modeling using 3DStudioMAX (Autodesk, 2008).

(b) Sound and Music Design using Soundforge (Sony Creative Software, 2008).

Figure 50: The Pre-Authoring Process.

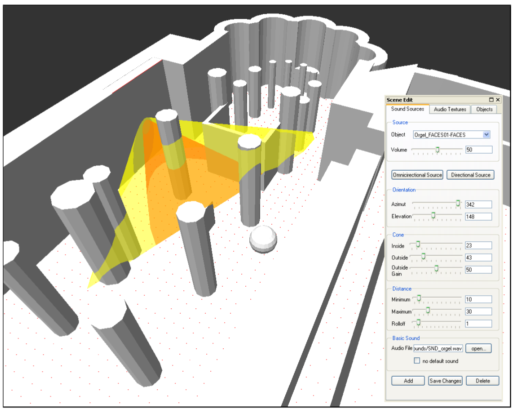
7.3.1 *Implementation*

The implementation of this authoring environment is performed using C ++ , and is based on the framework design as it was introduced in Section 5.5 and Section 6.3. The authoring environment therefore also utilizes OpenSG as scenegraph system and uses OpenAL/EFX for sound rendering and synthesis. OpenSG is used to load the 3D model as VRML data file and displays the 3D scene using an OpenGL visualization, compare with Figure 49. Individual 3D objects can be selected to assign and create 3D sound sources and/or auditory textures. The object data nodes are thereby extended by this information and saved to the VRML file after the authoring process is finished. The user interface for the authoring environment resembles a classic Windows GUI and was developed using the Qt library. The interface and 3D scene visualization is customizable through a selection of specific view screens and floating toolbars. These toolbars allow the creation, positioning and assignment of 3D sound sources and auditory textures to each virtual object or additional dummy objects, see also the following section.

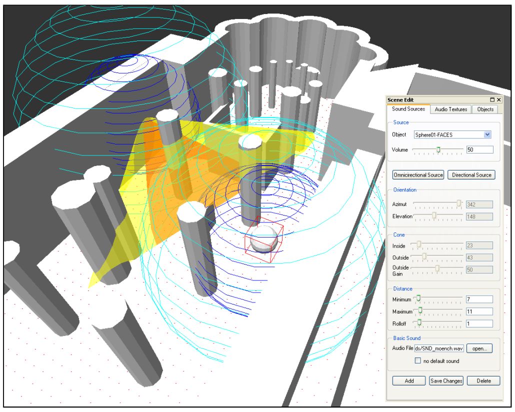
7.3.2 *Authoring Process*

As a conclusion for this chapter, an example is provided that shows the individual steps of the authoring process using the pipeline devised and the authoring environment developed. The example, as can be seen from Figure 49, is part of the authoring for an augmented audio reality application, which will be examined and evaluated in more detail in Section 9.5.2. The focus of this section is to discuss its design and construction and to illustrate the various aspects of the authoring process.

The design starts with the pre-authoring and the modeling of the 3D environment, see Figure 50a, as well as with a selection and mastering of suitable sound, speech and music samples, refer to Figure 50b. The result of this pre-authoring is a VRML data file that contains the geometry and topology of the 3D scene, as well as a pool of sound data for the sonification and interaction authoring. The VRML data is loaded into the authoring environment, compare with Figure 49, in which now the assignment of sound sources to certain scene objects, as well as the design of auditory textures starts.



(a) Creation of a Sound Source and Specification of the Source Direction.



(b) Specification of Distance Attenuation using two Spheres.

Figure 51: Sound Source Creation and Parameter Adjustments.

Sound Sources

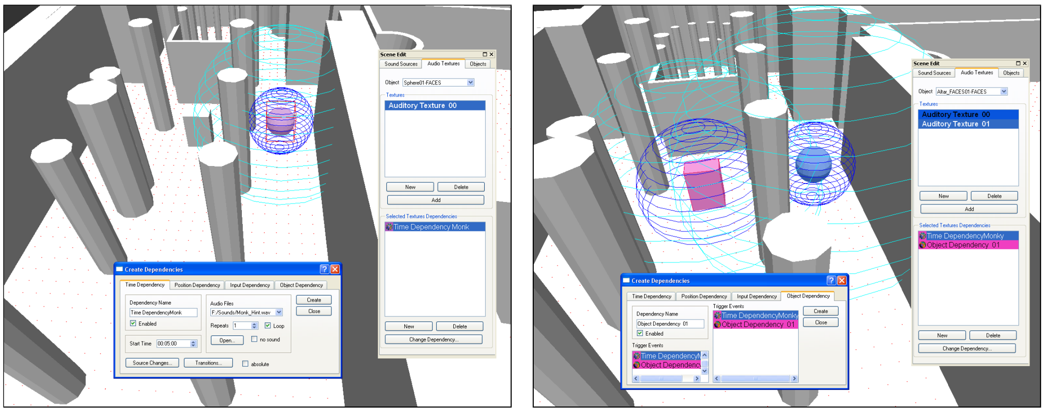
Authoring of Sound Sources.

The first part of the authoring is the creation of 3D sound sources and their assignment to individual scene objects. Figure 51 shows the creation of sound sources, as well as the setup of several source specific parameters, such as direction, distance attenuation, loudness and rolloff factor. As OpenAL is employed as underlying sound engine and used for the sound rendering and spatialization, the parameters and settings used are directly mapped to the OpenAL API (Hiebert, 2006). Additional source parameters are visualized using two semi-transparent cones, refer to Figure 51a, in which the transparency defines the loudness/gain and the orientation and size of the cone the direction and radiation pattern for this source. The application allows two different gains to be defined, one for the inner cone (red), and a second for the outer cone (yellow). A distance attenuation model describes the attenuation in between using the rolloff factor specified, compare with Figure 51b.

Auditory Textures

Authoring of Auditory Textures.

The definition and assignment of auditory textures is one of the most fundamental, but also one of the most flexible authoring steps, and is depicted in Figure 52. Figure 52a shows the definition of a time-dependency, while Figure 52b shows the dialog for authoring an object dependency. The sphere on the right hand side in Figure 52 is a so called dummy object, which serves as a placeholder and represents in this example the position of a virtual monk. This monk is represented acoustically through a snoring sound, because he is napping and the task of the player is to identify and activate this sound source later. The activation is triggered through proximity, in which an additional position dependency changes the state of the source (ie. wakes up the monk) and initiates a short dialog, refer also to Section 9.5.2. In this example, the time dependency is set to 5 minutes, and if activated, an auditory hint is played to guide the player in the direction of the monk, see also Figure 52a. This time dependency also triggers an activation of the altar object through a defined object dependency, refer to Figure 52b. Time and object dependencies can be employed in many different tasks and are a simple and intuitive way to integrate content into the 3D auditory environment and to narrate a storyline.



(a) Authoring of a Time Dependency for the Sphere Object (Monk).

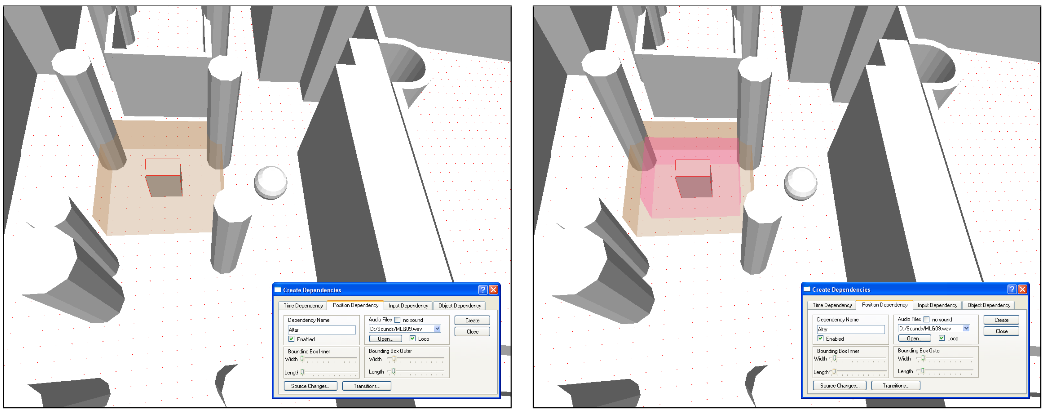
(b) Authoring of an Object Dependency between the Sphere (Monk) and Altar Objects.

Figure 52: Authoring of Auditory Textures.

Position Dependency

Authoring of a Position Dependency.

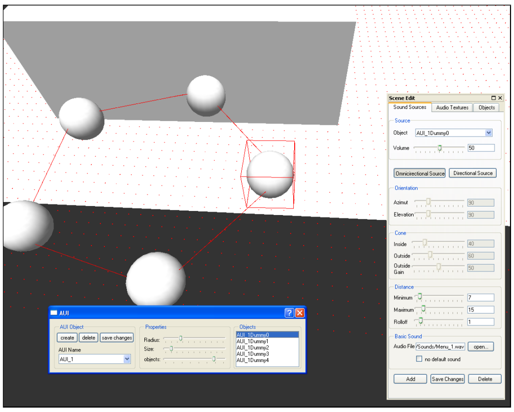
Another part of the auditory texture authoring is the creation of position dependencies, which are shown in Figure 53. Position dependencies are used as a proximity sensor to activate an object, and/or to trigger another object dependency. The authoring of position dependencies for the augmented audio example requires, due to the inaccuracies of the WiFi-based user positioning, the specification of an outer and an inner rim, refer to Figure 53. The inner rim is thereby used to activate an object, while the outer rim is used to deactivate it, see also the discussion in Section 6.2. The differences in length of both *boundary boxes* must be wide enough to accommodate the maximum error of the positioning system. Otherwise, a constant activation/deactivation of an object occurs, which might result in a choppy and very disturbing playback of sound. The same approach can also be used for the authoring of 3D virtual auditory environments, but as the avatar *positioning* is here controlled using a gamepad, the distance between the two bounding boxes can be minimized.



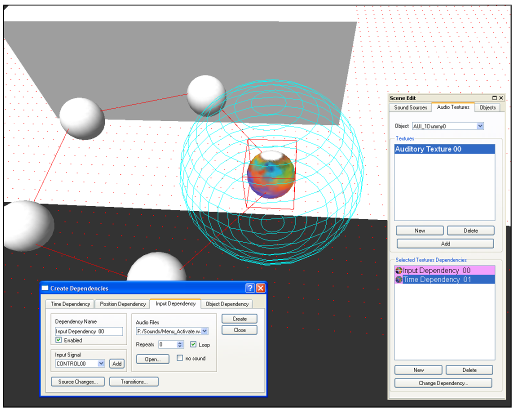
(a) Definition of the outer Rim (Source Deactivation).

(b) Definition of the inner Rim (Source Activation).

Figure 53: Authoring of Position Dependencies.



(a) Creation of a fivefold Ring Menu System.



(b) Definition of Auditory Textures with Interaction Dependencies and Assignment to the Menu Objects.

Figure 54: Authoring of a Ring Menu System.

Ring Menu System

The last example in Figure 54 shows the authoring of a ring-based menu system. Figure 54a shows here the creation of a five-fold circular menu system, for which one node is exemplary authored using an input dependency, see also Figure 54b. Each of these nodes is represented by a dummy object to which an auditory texture is assigned. This auditory texture is similar to the auditory textures discussed before, but can additionally handle an input dependency, eg. for user interaction. Upon interaction and an activation of a node, a certain action is executed. This action can be implemented using additional dependencies and also activate/change other scene objects using secondary dependencies. This section presented several of the previously discussed authoring techniques using an example from augmented audio reality. The authoring for other applications and tasks is very similar, and the main techniques and guidelines apply there as well.

Authoring of a Ring Menu System.

7.4 SUMMARY

This chapter discussed the requirements for an authoring and design of 3D virtual/augmented auditory environments. Several principles and design guidelines for the authoring of auditory environments have been discussed and were presented in the form of an authoring pipeline. Additionally, a 3D authoring environment has been devised and was implemented on top of the audio framework discussed in Chapter 5 and Chapter 6. This authoring environment allows the design of acoustically enriched 3D enhanced environments $\mathcal{E}$, and can be applied to a large variety of areas of application. Exemplarily, the design of an augmented audio reality scenario was presented and discussed.

This chapter basically concludes the discussion of 3D virtual/augmented auditory environments along their authoring and design. Later Chapter 9 returns to a discussion of 3D auditory environments with a presentation and evaluation of several applications and prototypic implementations. As an efficient auditory representation of these environments requires both, a realistic 3D sound spatialization and simulation of room acoustics, the following Chapter 8 is dedicated to a discussion of acoustic rendering techniques. Thereby it explores the possibilities of utilizing computer graphics rendering techniques and hardware for an effective and more realistic sound rendering and simulation.