INTERACTING WITH SOUND
AN INTERACTION PARADIGM FOR VIRTUAL AUDITORY WORLDS

Niklas Röber and Maic Masuch

Games Research Group
Institut for Simulation and Graphics
Magdeburg, Germany
nroeber|masuch@isg.cs.uni-magdeburg.de

ABSTRACT

The visual and the auditory field of perception react on different input signals from our environment. Thus, interacting with worlds solely trough sound is a very challenging task. This paper discusses methods and techniques for sonification and interaction in virtual auditory worlds. In particular, we describe auditory elements such as speech, sound and music and explain their application in diverse auditory situations, as well as interaction techniques for assisted sonification. The work is motivated by the development of a framework for the interactive exploration of auditory environments. The main focus for the design of this framework is the use in narrative environments for auditory games, communication purposes or general purpose auditory user interfaces.

1. INTRODUCTION

Vision is regarded as the most important of our senses from which we derive our environmental information, whereas audio plays a significant but often overlooked role in extending the horizon which is set by our visual system. With our vision apparatus, we are able to clearly identify the environment in front of us within the given viewing angle. However, as we can not see through (most) objects, we can only determine objects in our current field of view. Items that are hidden remain hidden. The advantage of the auditory field of view is that it enables us to additionally perceive sounds from behind and also from visually hidden objects. In combination with visual information, it draws a complete picture of our local environment, thus enabling us to proper interact with it [14].

In this paper we describe methods and techniques for sonification and interaction in virtual auditory worlds that provide the listener with enough information for clear interaction and navigation. The motivation behind this work is to create a complete catalogue of sonification and interaction techniques available for the exploration of virtual auditory spaces. The later discussed methods are the basis of a framework which will be used to evaluate the here discussed techniques and will serve as platform for experimenting with different setups for auditory user interfaces.

Audio in most applications is always used in combination with visual information, eg. movies, computer games or VR displays. Lately several applications have been developed which only utilize audio to convey information in so called auditory displays [8],[18] in which some utilize head-tracking functionality for easier interaction [12]. These displays communicate information only through auditory channels by using positional audio sources and by adding environmental reverberation information to these audio signals [10]. Several applications using auditory displays have been developed and evaluated, including audio only games [2] [6], auditory VR displays [8] and audio books [1]. So far, the largest field of application for auditory displays are assistive technologies for the blind and visually impaired [9], [26].

In the last decades important research has been accomplished in the field of audio based displays. Algorithms have been developed for the correct rendering of spatially localized sounds [15], [16], as well as for environmental audio simulations [11]. Auditory research has not focused on the technical side alone, the perception of spatialized and environmental sounds [22] or the design of sound sources and earcons are of equal importance [29]. Auditory displays are of major interest for assistive technologies, by guiding visually impaired people [13] or to assist fireman in smoked buildings [28]. Similar to the field of scientific visualization, sonification techniques have been employed to sequential data sets [25], [30] for analysis and exploration.

Several publications have studied sonification and interaction methods for auditory environments [7], but often focused on assistive technology rather exploring techniques for narrative environments or general interaction with virtual auditory spaces. Diverse approaches exist for sonification techniques in general auditory user interfaces [24], [19]. Wearable computer devices for augmented auditory reality were developed by Roy [23] and Sundareswaran [28] with the main focus on assisting in navigation and pathfinding in real environments.

The paper is organized as follows: In the next Section we describe auditory elements and the applicability for their usage in diverse auditory applications. Section three explains methods to convey non-auditory information within auditory channels and describes techniques for natural interaction within auditory environments. In the following, Section four presents the design of a framework, which is based on the methods discussed in the previous sections and implements the theory for evaluation. Finally Section five summarizes the work and states possibilities for future extensions.

2. AUDITORY ELEMENTS

In virtual worlds only a few specialized operations exist that can be performed without auditory or visual information present. Some of these audio only applications include audio books [1], audio games [2], sonification techniques for of scientific data sets [25] and assistive auditory displays for the navigation of visually impaired people [13], [9]. The task of these applications is different, and the usage of auditory elements to convey information to the listener is varying.

A close look at the foundation of audio reveals three auditory elements, of which every auditory composition in our environment is composed off:

- Speech,
- Music, and
- Natural or artificial sounds.

The auditory spectrum is composed of signals, which itself are constructed by these three elements. Speech is an oral transmission of information by using words as an abstract representation and mainly used for communication. Music is the concatenation of tones, resulting in harmonic compositions and often evolves emotions. Music is generally used on top of speech and sound to accompany the presented information. The largest group is build by natural and artificial sounds, which describe audio signals that depict a physical object or process, eg. starting a cars engine or the sound of leaves in a tree. To anticipate sonification, each of these elements is assigned a piece of information to be conveyed. Depending on the role of the information, sounds occurring in a scene are grouped as main sounds, supporting sounds and ambient information.

Each of these auditory elements is suited best to express a certain piece of information. In general, speech is mainly used to transmit knowledge, news or advice, like the oral description of a scene. Many applications use speech as narrative element, but the main purpose of speech is the communication between two or more characters. Music is often used to enrich the sound field and offers a deeper immersion into the activity. Music is especially very well suited to influence the listener and to change his mood. Faster music leads to a more aggressive behaviour, like in racing games, as slow music is often used to chill. The pitch of music can be used to describe either a dangerous place, dull music, or a peaceful scene using light music. Sounds have the largest variety of the three elements, ranging from *describing sounds*, which provide information about the local environment, to *artificial sounds*, which can be employed in virtual environments as agreed signals to identify objects, see also Section 3.1. Complex sounds, like the interaction between objects and the resultant difference in sound, are described through auditory textures which specify the sound pool available for these objects.

As discussed earlier, several applications have been developed that utilize audio signals as the main transmitter for information. The interactivity in these applications ranges from passive listening (audio books) to highly interactive programs (computer games). Most of them use all three basic auditory elements to commute information to the listener. In the following, a short survey of how the auditory elements are used by these applications is provided.

Several auditory assistive devices have been developed to aid in navigation and orientation for the blind or visually impaired. Many of these auditory displays use speech or guiding tones [9]. The latest development are assistive devices that use both positional sound and sonification to augment the auditory environment [13]. The interaction with these devices often employs GPS for location determination and head-tracking for view specification [26]. Depending on these input information the location is looked up in a digital map and described acoustically through sound and speech. The drawback to most of these systems is that they can not react on current changes, like construction work.

Scientific sonification is the analysis and exploration of abstract data through non-speech sound. The sonification of sequential or time-varying data is often faster than the graphical visualization and allows an easier perception. Examples for sonification are applications in seismic studies [25] or the Geiger-Müller radiation counter. Sonification employs only sounds and tones. Using the segregatable attributes of sound, like pitch, timbre, loudness, spatial location and the organization over time, many data sets can be monitored acoustically.

Audio books are auditory presented books where the story is told by a narrator. The audio field is enriched by sounds describing the currently depicted scene. An audio book can be interpreted as an auditory scene, where the listener is the passive observer that is guided through the story by the narrator. The majority of information is presented orally through speech. Here speech is used for the narrator as well as for the communication between the different characters occurring in the books story. The narrator provides the listener with an initial setup of the scene and during the advance of the story with updates on important changes. Music and describing sounds can be used to enrich the scene and to allow a better immersion into the story. Here, music is responsible for the atmosphere and to influence the listeners feelings. Sounds from objects within the scene allow for a better interpretation of the environment and to observe activities that are currently not in the listeners focus. Audio books are commonly presented in stereo sound. No spatial information of the auditory scene is involved and generally no interaction is possible.

Some audio only games have been developed in the recent years [5], several of them as hybrid applications which on request allow a visual display of the scene [6]. The variety of genres ranges from action games, like car racing [4], to interactive stories [5]. The interaction with most of these games is similar to other computer applications, by simply using a keyboard. For sonification, sounds are employed that provide the player with additional information about the surroundings. In interactive stories speech is used as narrative element and to describe the local environment. Other games use speech to warn of obstacles or to provide information for possible interactions with objects. Music is used rarely to not clutter the auditory display with too much information. Audio games employ both, positional audio and environmental sound sources.

In conclusion, nearly each application employs the auditory elements in a different manner. Few of them use all auditory elements, some only one or two, depending on the applications task. In scientific sonification it is the analysis and the exploration of the data, for audio books the story and for audio games the game status and current perspectives. The main difficulty is the correct conveyance of the underlying information. The next two sections focus on exact these problems with summarizing methods and ideas for sonification and interaction in virtual auditory worlds.

3. SONIFICATION AND INTERACTION

The most crucial part in audio only applications is the correct transmission of non-auditory information through auditory channels. It is nearly impossible to describe an image using non-speech sound or to visualize an opera or a song using a picture or an animation. As a result of this, speech has emerged. The ancient saying *A Picture is worth a thousand words.* emphasizes this difficult approach even by using an verbal description. In either ways, several examples exist, some of them with pretty good results. One of the earliest examples is the classical masterpiece of “Peter and the Wolf” by Prokofjew [20] which tells a story in which each character is assigned an instrument. The interaction between them creates wonderful music. Another example is "Pictures at an Exhibition" from Mussorgski and Ravel [17], [21] and later an electronic version from Tomita [27]. Both describe the experiences and the relations of the composer to the paintings of the Russian artist Viktor Hartmann.

The next two sections investigate methods to convey non-auditory information by only using auditory channels and how to properly interact with virtual auditory environments. The focus is on the process of hearing and hearing behaviours and how they can be used for natural and easy interaction. Both sections are centred around sonification and interaction with virtual auditory worlds. Some of these techniques have been described before, but the focus here is on how these techniques can be employed for narrative games taking place in 3D auditory environments.

3.1. Sonification

Sonification is defined as the mapping of abstract data to non-speech sound and used to transmit arbitrary information through auditory channels. We constantly perceive from our environment information through our sensory apparatus, mostly vision and sound, which is analyzed and interpreted. As we receive large amounts of data, some is filtered out at early stages and not actively perceived. Strong correlated to sonification is *interaction*. For every interaction with the environment there is a reaction in sonification which provides the listener (interactor) with feedback information, see also Section 3.2.

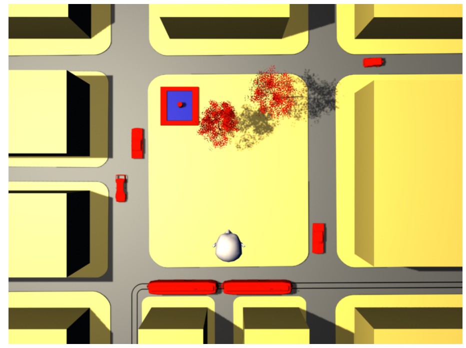
With the focus on sonification of auditory worlds, several *information groups* can be identified. These groups can be summarized by the following questions:

- Where is something?
- What is this?
- What can I do with it?

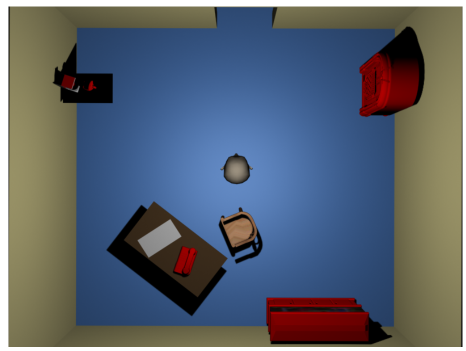
The fist question deals with environmental information that allows for orientation and navigation within the world. The second question characterizes the information which is necessary to identify and analyze objects, while the last question states possible interactions with interactable objects.

Important here is the first group from which we receive information about the topology of the environment. Visually, this is done by the viewing system that determines shape, size, texture and the location of objects. These properties additionally often allow to conclude about the objects function. For auditory spaces, this perception is a bit more complicated. Although, we are able to hear spatialized sound and can compare the distances between several sound signals, these audio signals have to be sequential and can only be perceived over time. Important for a sound sonification is that the listener never reaches a silent spot, which means no information, nor that too many audio signals clutter the display and complicate further perception.

For navigation and orientation within an auditory world, several methods can be employed. Navigation strongly depends on the interaction with the environment, see also Section 3.2. To determine the own position one must be able to identify the local environment. This can be done by using auditory landmarks. Auditory landmarks are distinct prominent sound sources, spread over the environment which assist the user in orientation. Figure 1 shows a graphical illustration of auditory landmarks in two different scales: on environmental (city) and room level. For local orientation, the sounds emitted by objects in a close vicinity (red) help to identify the environment and allow for navigation.



(a) City level.



(b) Room level.

Figure 1: Auditory Landmarks.

All objects in an auditory environment must emit audio signals, otherwise there are inaudible and not perceivable. Thus, every object in the environment is interactable. Depending on the type of interaction, interactable objects can be grouped like:

- Obstacles,
- Portals, and
- Interactables.

The first group is defined by objects that actively shape the environment and which are basically not interactable, except that they are barriers which can interfere in the free exploration of the environment. The only interaction possible is collision and obstruction. Object bound sounds can be employed to sense the obstructions which increase in volume when the barrier gets closer. The second group also specifies parts of the environment as these classifies objects we can actively interact with to change our position in the environment, like doors, escalators, stairs or teleporters. This includes the ground floor, whose auditory texture provides information about the underground material. The interaction with some of these objects might be obvious, but in an audio only environment, one must perceive information that the elevation changes (climbing stairs) or that one passed trough a door into a new room. Here transitional sounds can be used which announce the transit through a portal. For stairs, escalators or elevators it is helpful to verbally notify when a new level is reached.

The third group is represented by objects that really are interactable. These are usually smaller objects, like buttons, a radio or a telephone. While we are familiar with most of these objects functionality, for some objects this has to be explained, either verbally or through describing sounds. For an example, the functionality that a button can be pressed on a radio is sonified by a clicking sound, which can be enhanced by a verbal description. For most objects, if the functionality is known, the possible interactions become clear. More difficult is the interaction with objects unfamiliar to the listener. Here, on a first encounter most efficient is a verbal description of the object itself and its functionality and the introduction of describing sounds for later recognition. After one successful interaction, the verbal description does not has to be used again, but can be evoked manually. When interacting with objects, its is very important to receive feedback information about the success of the current operation. This is done by changing the describing sound for this object, or by changing the environmental information, for instance, when the interaction was to ignite an engine. Focusing on a specific object for either interaction or exploration means to highlight this object. This can be done by suppressing all environmental sounds, except for this object and enabling additional audio cues, like oral descriptions or describing sounds.

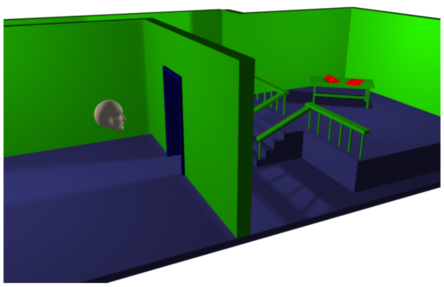


Figure 2: The three groups of interaction

A descriptive illustration for the different interaction groups can be seen in Figure 2, where the wall and the handrail define obstacles (green), the door and the stairs portals (blue) and finally the paper and the telephone on the table are interactables (red).

Difficult in auditory environments is to sonify the change in light intensity. As we perceive only auditory information, it actually does not matter if the scene is lit at all or completely dark. But incorporating lighting conditions is a strong dramaturgical element and can be employed in auditory environments as well. In movies or radio plays, when the hero reaches a dangerous place, dark and dull music is played to enhance the feeling of danger and mystery. Opposite, to sonify a bright and safe environment, very light and crystal music is utilized. Additionally, all sounds from the environment, this excludes the narrator, can be changed in pitch to reflect this manner.

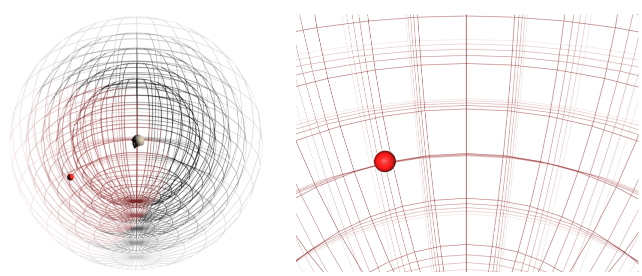
Global changes within the environment, changes in time or changes in the story, should be announced by the narrator. Communication with other characters, real or artificial, are done by simply talking via a microphone. Each character is assigned a personal sound ID that is used for communication requests.

In summary, the sonification of auditory environments implies several different problems. With the shown techniques, all necessary tasks to gain information from a virtual environment can be accomplished. Methods for navigation, object exploration and interaction have been discussed. Sonification strongly dependents on interaction, and vice versa. Therefore, these techniques shall be discussed next.

3.2. Interaction

To effectively convey information to the listener, the hearing behaviour has to be incorporated into the sonification process. Naturally, humans slightly tilt their head to determine the location of a sound source in ambiguous cases or when listening very precisely. Technically, this behaviour results in a different input angle for the signal and helps to accurately locate the sound source and additionally to listen more focused. When interacting in virtual auditory environments, special care has to be taken for the sound localization process as this is mandatory to determine the own position within the environment. Additionally, spatialized sound can assist in the process of discrimination between several audio signals if they originate from different locations. However, if too many sound sources are presented at the same time, the auditory display can easily get cluttered. In this section we discuss techniques that should be used for proper interaction within auditory spaces and which support the sonification process, see also Section 3.1.

As with sonification, also interaction can be split down to navigation, interaction with objects and communication with other characters. For the navigational part, the user must be able to control his position and, with the help of the sonification techniques discussed above, receive immediate feedback which enable him to evaluate wether the last move went into the right direction or not. To move around, either a keyboard or a joystick are most common. The point and click technique known from several games can be employed as well, but is more complicated to use, as the player has to concentrate on the environment as well as on the auditory cursor to place the character for the next move. A joystick in combination with force feedback could enable haptic input and provide the player with additional information and guidance. When further proceeding in a certain direction is not possible, force feedback can tell the user by denying this action. Without haptic feedback, this can be implemented using sonification techniques where the player receives acoustical feedback about obstacles, see Section 3.1. Secondary, a specially designed joystick is more suitable for mobile interaction, providing easier navigation, whereas the input using a keyboard is more restricted.



(a) Overlook. (b) Listeners view.

Figure 3: Auditory Cursor.

The methods for interacting with objects depend on which interactions are possible. A special class are the techniques to interact with the environment to receive information about the local surroundings. With careful listening and by using trackable headphones one is able to precisely determine the position a sound signal originates. If the user turns the head, the scene *rotates* in the opposite direction. This head-tracking [12] provides an immediate response from the system on the changed input conditions and offers a deeper immersion into the auditory space. Sometimes, the information provided by this technique is not enough and the object type (sound source) or object interaction is unclear. Here, further interaction with the scene can be made possible by using either an auditory cursor or a flashlight like radar device.

In contrast to a regular computer cursor, the auditory cursor is able to work within a complete 3D environment. The basis is a spherical coordinate system which also specifies the grid (Fig. 3). As real 3D interaction is possible, the grid can be thought of several spherical shells layered into one another. The auditory cursor is a special sound object which can be placed on this grid and move along the grid intersections (red dot in Fig. 3). The location on the sphere is clearly audible using positional sound techniques. The depth is encoded using the pitch and the loudness of the sound signal. A more dull sound is farther away than a clear sound. The cursor snaps to the grid and for more precise placements, oral notifications of the coordinates can be invoked. Figure 3(a) shows five spheres and the listeners head, where the viewing direction is visualized through a red light. Figure 3(b) shows the scene through the *eyes* of the listener.

Figure 4: Radar Interactor

The radar device is another interaction technique which allows for a precise scanning of the scene and high quality information. The radar can be thought of as a flashlight, which is used to illuminate the scene. Everything that gets *lit* by the radar answers in a predefined manner, by either a verbal description or an agreed sound signal, see Figure for a descriptive illustration 4. Every response is amplified by the distance, so that closer objects appear closer. The radar can also be employed as a sonar to assist in the detection of certain items. The iterating signal changes in loudness and frequency if the object moves into focus and gets closer. Using an interaction button on the radar, objects can be selected for further investigation. The interaction with objects, like turning on the radio, can be accomplished by using either the radar or the joystick. The object in focus becomes highlighted through the previously discussed sonification techniques and can now be used for interaction.

The interaction auditory displays are best suited for is communication. This (oral) information exchange is the fastest way to interact with artificial avatars in a scene, or other characters in a multi-player environment. As the avatars are objects from within the environment, the communication is positional and the speech originates from the characters location in the scene. The communication itself can be implemented straightforwardly, but care has to be taken for the communication with artificial characters. The speech has to be parsed, analyzed and a proper answer found which itself is presented orally. Speech analysis is difficult and restricted to a predefined catalogue of words. Similar applications are the old day text adventures, where the communicative media was text.

Besides sonification, interaction is one key element for a deep immersion into the experienced environment. Several techniques have been discussed in the last section, emphasizing on interaction techniques that aid in orientation and navigation and also for close object operations. For the success of these methods, the most important part is that these methods can be implemented with realtime feedback. If the latency is too large, the interaction and the resultant sonification will be perceived disjunct.

4. DESIGNING A FRAMEWORK

Based on the discussion in the previous sections, this section layouts an initial design for a framework, which allows an intuitive and easy interaction with narrative environments. The focus for this framework is the later use in narrative environments for interactive adventures by utilizing only positional and environmental audio as information sources. The goal is further to design this framework as open as possible to allow an easy adaptation to other fields, and explore possible applications in tele-conferencing, audio-action games, mobile auditory displays or general non-visual user interfaces based on 3D audio. This section discusses work in progress.

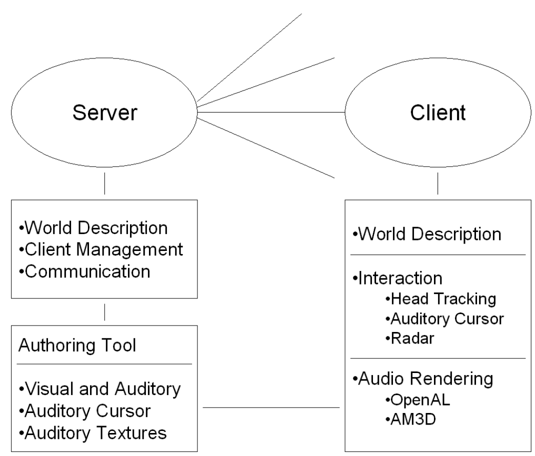


Figure 5: Framework Layout

The motivation behind this work is to create an immersive non-visual user interface which is able to interactively guide a user through either entertaining auditory worlds or to use these interfaces for mobile applications where the desire for a free view is mandatory. The design was mainly motivated by the actions the listener should be able to accomplish. Most of these methods for sonification and interaction discussed earlier are now being implemented and will be evaluated within this framework. Tabular 1 shows an overview of the desired actions and their appendant methods for efficient sonification and interaction.

User Action	Sonification	Interaction
Determine Position	landmarking, object sounds, oral scene description	head tracking
Navigation (without moving)	landmarking, object sounds	head tracking, radar
Navigation (with moving)	landmarking, object sounds	head tracking, joystick
Object Analysis	object sounds, oral object description	radar
Object Interaction	object sounds	joystick, radar
Character Interaction/Communication	object sounds	microphone

Table 1: Overview of User Actions in an auditory environment

An overview of the frameworks architecture can be seen in Figure 1. The layout was inspired by general game engine design issues [31], [3]. The framework is build as a client-server architecture where both, client and server administrate a description of the virtual world. The server manages all characters (clients) and handles the communication between them. The clients are independent systems which can also work without the server. Every client has a description of the world and renders the auditory scene from its own perspective. Depending on the audio hardware available, the audio rendering uses currently OpenAL, and will later be extended to DirectSound3D and the commercial library available from AM3D. The implementation itself is entirely in C++ and for the world description XML is employed. The auditory environment of the virtual world is authored using a visual control renderer. As for audio rendering limitations, the current implementation platform is Windows, but a cross-platform implementation is targeted.

5. CONCLUSIONS AND FUTURE DEVELOPMENT

The knowledge of which methods are applicable for the sonification of specific data or information is very important. Sonification and interaction both play an important role in the perception of auditory worlds, and have to be considered together to emphasize the result. Several techniques have been discussed, which are also applicable in other fields of auditory display.

As shown in Section two, all audio signals can be split up in three large groups, namely speech, music and sound. These auditory elements are used differently in each single application, depending on the sonification goal. Then we discussed methods and techniques for sonification and interaction within virtual auditory worlds. We claim that the interaction techniques have to be as natural as possible to closely mimic the process of natural hearing, which results in a deeper immersion into the virtual environment and a better perception of the conveyed information. This is a necessity for being able to correctly decode the audio signals through sonification in order to benefit from all advantages through auditory perception.

The framework discussed in Section 4 is work in progress. Unlike other audio engines for computer games, we explicitly focus on sonification and interaction techniques for audio-only games, like auditory adventures. Several prototypes exist and the first experiments are promising. The next steps include the design and a practical evaluation of these techniques using a narrative auditory scene. For most current information, we refer to our website[1].

Another challenge is the design of auditory spaces. Here a more advanced authoring tool is under development which allows both, a visual and an auditory-only interface to design such worlds. *Auditory Textures* are used to model the complete behaviour of surfaces under different conditions, like a dry, rainy or wet street surface. Additionally, for the auditory modelling, a spherical grid is used along with an auditory cursor and the previously discussed sonification techniques to be able to model and create the auditory environment completely non-visual through audio only.

[1] `http:\\isgwww.cs.uni-magdeburg.de/~nroeber/research/audio/index.html`

6. REFERENCES

[1] http://www.audiobooks.com. Website, 2003.

[2] http://www.audiogames.net. Website, 2003.

[3] James Boer. *Game Programming Audio*. Charles River Media, Inc., 2002.

[4] Developed by Christopher Lewis. Audio formula 1. Published by Christopher Lewis, 2003. PC.

[5] Developed by Fachhochschule Stuttgart (Tanja Dannecker, Matthias Pasedag, Christa Stoll, and Heinrich Sturm). Der Tag wird zur Nacht, 2003. PC.

[6] Developed by Pin Interactive. Terraformers. Published by Pin Interactive, 2003. PC.

[7] Kai Crispien and Klaus Fellbaum. A 3D-Auditory Environment for Hierarchical Navigation in Non-visual Interaction. In *ICAD*, 1996.

[8] Gerhard Eckel. A Spatial Auditory Display for the CyberStage. In *ICAD*, 1998.

[9] C. Frauenberger and M. Noisternig. 3d audio interfaces for the blind. In *ICAD*, 2003.

[10] Thomas Funkhouser, Jean Marc Jot, and Nicolas Tsingos. Siggraph Tutorial 45: Sounds Good to Me! Computational Sound for Graphics, VR, and Interactive Systems. In *Siggraph*, 2002.

[11] Thomas Funkhouser, Nicolas Tsingos, Ingrid Carlbom, Gary Elko, Mohanand Sondhi, and James West. Modeling Sound Reflection and Diffraction in Architectural Environments with Beam Tracing. In *Forum Acusticum*, 2002.

[12] Lalya Gaye. A Flexible 3D Sound System for Interactive Applications. In *CHI*, 2002.

[13] Robert W. Massof. Auditory assistive Devices for the Blind. In *ICAD*, pages 271–275, 2003.

[14] Margaret W. Matlin. *Sensation and Perception*. Pearson Allyn & Bacon, 1987.

[15] M. Morimoto and Y. Ando. On the Simulation of Sound Localization. pages 167–174, 1980.

[16] Masayuki Morimoto, Kazuhiro Iida, and Motokuni Itoh. 3-D Sound Image Localization by Interaural Differences and The Median Plane HRTF. In *ICAD*, 2002.

[17] Modest Mussorgski. Pictures at an exhibition, 1874.

[18] M. Noisternig, T. Musil, A. Sontacchi, and R. Höldrich. A 3D real time Rendering Engine for binaural Sound Reproduction. In Eoin Brazil and Barbara Shinn-Cunningham, editors, *ICAD*, volume 9, pages 107–110, 2003.

[19] Ian Pitt and Alistair Edwards. *Design of Speech-based Devices*. Springer, 2003.

[20] Sergej Prokofjew. Peter and the wolf, 1936. Op 67.

[21] Maurice Revel. Pictures at an exhibition, 1922.

[22] Marie Rivenez, Carolyn Drake, Anne Guillaume, and Sebastien Detry. Listening to Environmental Scenes in Real Time. In *ICAD*, 2002.

[23] Deb Roy, Nitin Sawhney, Chris Schmandt, and Alex Pentland. Wearable Audio Computing: A Survey of Interaction Techniques. 1997.

[24] Nitin Sawhney and Arthur Murphy. ESPACE 2: An Experimental HyperAudio Environment. In *CHI*, pages 105–106, 1996.

[25] S. D. Speeth. Seismometer Sounds. In *Journal of the Acoustical Society of America*, volume 33, pages 909–916, 1961.

[26] Thomas Strothotte, Helen Petrie, Valerie Johnson, and Lars Reichert. Mobic: an aid to increase the independent mobility of blind and elderly travellers. In *2nd TIDE Congress*, 1995.

[27] Isao Tomita. Pictures at an exhibition, 1975.

[28] Steven Chen Reinhold Behringer Joshua McGee Clement Tam Pavel Zahorik V. Sundareswaran, Kenneth Wang. 3D Audio Augmented Reality: Implementation and Experiments. In *Second IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR 03)*, 2003.

[29] M.L.M. Vargas and S. Anderson. Combining Speech and Earcons to assist Menu Navigation. In *ICAD*, 2002.

[30] B.N. Walker and J.T. Cothran. Sonification Sandbox: A graphical Toolkit for auditory Graphs. In *ICAD*, 2003.

[31] James Watt and Fabio Policarpos. *3D Games: Real-Time Rendering and Software Technology*. Addision Wesley, 2002.