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 respond 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, it describes auditory elements such as speech, sound and music and discusses 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 which will be used to evaluate the later discussed techniques. The main focus for the design of this framework is the use in narrative environments for auditory games, but also for general purpose auditory user interfaces and communication processes.

1. INTRODUCTION

Vision is considered as being the most important of our senses from which we derive the majority of our environmental information. Hearing is often assumed to play a minor role only, but it assists in extending the horizon that is set by our visual system. Using visual information, we are able to clearly identify the environment in front of us within the defined viewing angle. However, as we can not see through most objects, we can only identify objects that are visible or visible through reflection or refraction. One advantage of the auditory field of view is that it enables us to perceive audible information from objects that are hidden or outside the viewing cone. Combined with visual information, it draws a complete picture of our local environment, thus enabling us to proper interact with it [18].

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

Auditory information is often used in conjunction with visual information to assist in the perception of the depicted information. Applications for 3-dimensional sound exist in various fields, ranging from VR-displays [9],[22], auditory games [2] [15], radio plays and audio books [1] to tools to aid in the navigation for the visually impaired [10], [29]. A good introduction into 3-dimensional and environmental sound synthesis can be found in the books by Begault [4] and Garas [13]. Many auditory applications have developed specific sonification and interaction techniques. Some utilize head-tracking functionality to intensify the immersion and to allow for an easier, more intuitive interaction with the auditory space [14],[11]. Various studies exist that evaluate the quality of the sound perception [19], [20],[12] and the suitability of sonification and interaction methods [26],[32]. Several publications have studied methods for sonification and interaction within auditory environments [7], but often focused on assistive technologies rather exploring the potentials for narrative environments, as found in auditory adventure games. Wearable computer devices for augmented auditory reality were developed by Roy [27] and Sundareswaran [30] with the main focus on assisting in navigation and pathfinding in real environments. Similar to the field of scientific visualization, sonification techniques have been employed to sequential data sets [28], [33] for analysis and exploration.

The paper is organized as follows: In the next Section we describe auditory elements and their applicability in various auditory situations. Section three discusses methods to convey non-auditory information within auditory channels and describes techniques for intuitive interaction with auditory spaces. In the following, Section four presents the design of a framework that is based on the methods discussed earlier and implements the theory for evaluation. Finally Section five summarizes the work and states possibilities for future explorations.

2. AUDITORY ELEMENTS

All operations performed in virtual environments require either visual or auditory information to be present. Some applications have been developed that can substitute the visual knowledge through auditory descriptions. Examples include audio books and radio plays [1], audio only games [2], sonification techniques for scientific data sets [28] and assistive auditory displays for the navigation of visually impaired people [17], [10]. The pattern of the auditory signals used and the functionality vary depending upon the applications requirements.

Using physical explanations, sound is a mechanical vibration which is transmitted by an elastic medium. It can also be described as being the audible part of a transmitted signal that was emit by a physical process. On a larger level, these auditory phenomena can be grouped together, revealing three basic auditory elements of which every audible sequence in our environment is composed off:

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

The auditory spectrum is composed of auditory sequences, which itself are constructed by these three elements. Speech is a verbal transmission of information by using words as an abstract representation and is mainly used for communication. Music is the concatenation of tones, resulting in harmonic compositions and often used to express or trigger 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 rustling in a tree. Technically, music and speech are special cases of sound. Depending on the type and the importance of information, sounds can further be 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 communication between two or more characters. Music is used to enrich the sound field and often combines it with an emotional note. Therefore, it is especially very well suited to influence the listener and to affect his mood. It is well known that faster music leads to a more aggressive behaviour, like in racing games, as slow music, which is often used to chill. The pitch can be used to control the darkness of a scene where dull music is used to describe a darker area, while light music is utilized for bright and peaceful scenes. Sounds have the largest variety of the three elements, ranging from *describing sounds*, which provide information about the local environment, to *artificial sounds*, that can be employed in virtual environments as agreed signals to identify objects, see also Section 3.1. Complex sounds, like the interaction between several objects and the resultant differences in sound, can be 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). Depending on the applications task, some of these elements are used with different intentions, which shall be discussed next.

Auditory assistive devices have been developed to aid in the navigation and orientation for the blind or visually impaired. Many of these auditory displays employ speech or guiding tones [10] to provide directions. The latest development are assistive devices that use both, positional sound and sonification to augment the auditory environment [17]. The interaction with these devices often employs GPS sensors for location determination and head-tracking for view specification [29]. Depending on these input information, the location is looked up in a digital map and described acoustically through sound and speech. The drawback of most systems is that they can not provide current information, like construction work on a sidewalk.

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 [28] and the Geiger-Müller radiation counter. Sonification usually only employs 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 and radio plays are auditory presented books or plays 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 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 enhance the scene and assist in a deep immersion into the story. Here, music is mostly responsible for creating the atmosphere and to influence the listeners feelings. Sounds from objects within the scene allow a better interpretation of the environment and to observe activities that are currently not in the listeners focus. Audio books and radio plays 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 [8]. Several are created as hybrid applications, which on request allow a visual display of the scene [15]. The variety of genres ranges from action games, like car racing [16], to interactive stories [8]. The interaction with most of these games is similar to other computer applications, by simply using a keyboard for input. For scene 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 employ speech to warn of obstacles, or to provide information for possible interaction with objects. Music is used rarely, to not clutter the auditory display with too much information. Audio games often employ both, positional audio and environmental sound sources to accurately render the scene.

In conclusion, nearly each application employs the auditory elements in a different manner. Few of them use all auditory elements, some only employ one or two depending on the applications requirement. In scientific sonification this is the analysis and the exploration of the data, for radio plays the presentation of the story and for audio games the display of the game status and the 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 using a picture or an animation. As a result, speech has emerged to assist in the communication process. In both ways, several attempts exist, some of them with pretty good results. One of the earliest is the classical masterpiece of “Peter and the Wolf” by Prokofjew [24] 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 [21], [25] and later an electronic version from Tomita [31]. 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 behaviour and how they can be utilized for natural and easy interactions. Both sections are centered around sonification and interaction with virtual auditory worlds. Some of the techniques have been described before, but the focus is on how these techniques can be employed for narrative environments, like computer games taking place in a 3D auditory world.

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 information from our environment through our sensory apparatus, mostly vision and sound, which is filtered, analyzed and interpreted. As we receive large amounts of data, some is filtered out at early stages and not actively perceived. Strongly correlated to sonification is *interaction*. For every interaction with the environment there are changes in sonification that provide the listener (interactor) with feedback information, see also Section 3.2.

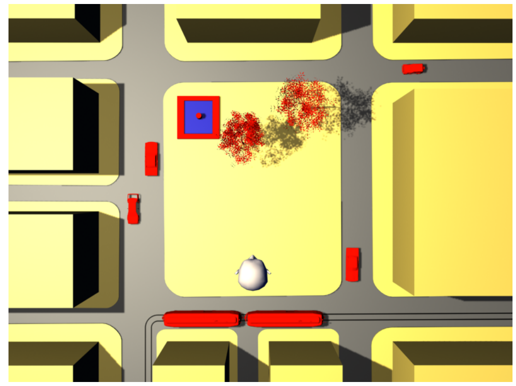
With the focus on sonification of auditory worlds, several *information groups* can be identified. These groups can be summarized by the following questions:

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

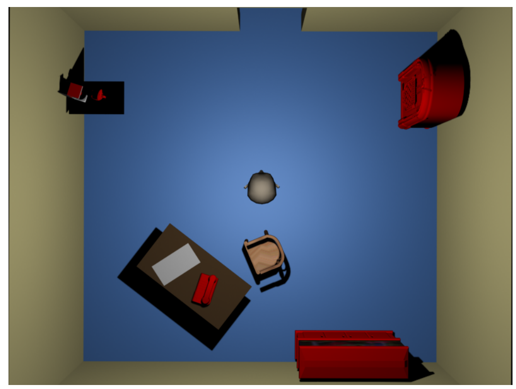
The first question deals with environmental information that allows for orientation and navigation within the world. The second question characterizes the information that is necessary to identify and analyze objects, while the last question states possible interactions with interactable objects.

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 often allow to conclude about the objects functionality. 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 sources, these audio signals have to be sequential and can only be perceived over time. However, it is easily possible to monitor several different sound sources at the same time [4],[6]. Important for 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 the information 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



(a) City level.



(b) Room level.

Figure 1: Auditory Landmarks.

environment. This can be done by using auditory landmarks. Auditory landmarks are distinct prominent sound sources, spread over the environment that 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 proper navigation.

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 as:

- 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 that 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 these obstructions with an 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 tele-porters. 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 another room. Here transitional sounds can be employed that 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 of them 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 with 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, it 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 a cars 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 that reveal more detail information about this object and its surroundings.

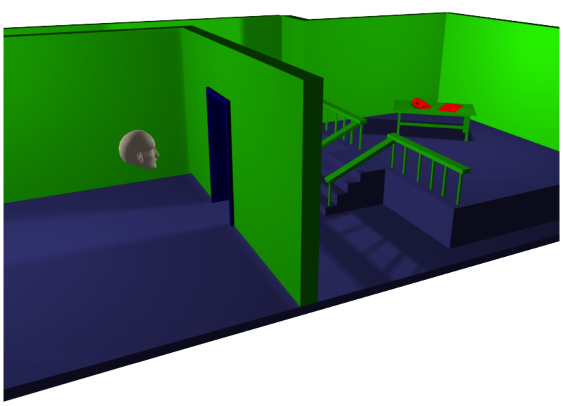


Figure 2: The three groups of interaction

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

More difficult is the sonification of 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 should be used for auditory scenes as well [3]. 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 headset. Each character is assigned a personal sound ID that is used for communication requests.

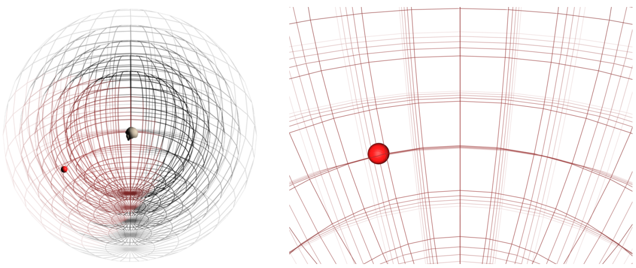
In summary, the sonification of auditory environments comprises several challenges. With the shown techniques, all necessary tasks to gain information from a virtual environment can be accomplished. Methods for navigation, object exploration and interaction have been discussed. Sonification strongly depends on interaction, and vice versa. The next section focuses on interaction techniques in more detail.

3.2. Interaction

To effectively convey information to the listener, the hearing behaviour has to be incorporated into the sonification process. Humans often 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 sound 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, resulting in a meaningless sonification. In this section we discuss techniques that can be used for proper interaction within auditory spaces and which support the sonification process, see also Section 3.1.

As with sonification, interaction can be split down to navigation, the 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 that enables 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 appropriate. 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 an auditory cursor to correctly 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 notify the user by denying this action. Without haptic feedback, this can be implemented using sonification techniques where the player receives acoustical feedback about obstacles and occlusions, see Section 3.1. Secondary, a specially designed joystick is more suitable for mobile interaction and provides an easier navigation, whereas the input using a keyboard is more restrictive.

The methods for interacting with objects depend on which interactions are possible. A special class is build by 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 provides an immediate response from the system on the changed input conditions and offers a deeper immersion into the auditory space [14]. Sometimes, the information provided by this technique is not sufficient 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.



(a) Overlook. (b) Listeners view.

Figure 3: Auditory Cursor.

Opposite to a general 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 cursors location on the sphere is clearly audible using positional sound techniques. The depth is encoded using the pitch and the loudness of the cursors sound. A more dull sound is farther away than a clear sound. The cursor snaps to the grid and for more precise placements, verbal notifications of the coordinates can be invoked. Figure 3(a) shows five spheres and the listeners head in the center. The viewing direction is visualized through a red spotlight. 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 a precise scanning of the scene and high quality information feedback. 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 4 for a descriptive illustration. Every response is amplified by the distance in a way 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 a joystick. The object in focus becomes highlighted through the previously discussed sonification techniques and can now be used for interaction.

A very natural interaction with auditory displays is communication through speech. This verbal 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 [23]. Similar applications are the old day text adventures, where the communicative media is 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 as being 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 and general non-visual user interfaces based on 3D audio. This section discusses work in progress.

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.

An overview of the frameworks architecture can be seen in Figure 1. The layout was inspired by general game and sound engine design issues [34], [5]. 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 that can also be used without being connected to the server. Every client has a description of the

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

world and renders the auditory scene from its own perspective. Depending on the audio hardware available, the 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 XML, which is used for the world description. The auditory environment of the virtual world is authored using a visual control renderer. As for current audio rendering limitations, the present implementation platform is Windows, but a cross-platform implementation is targeted.

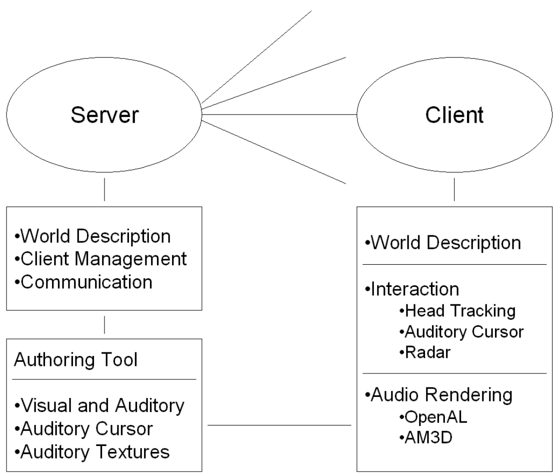


Figure 5: Framework Layout

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 to other fields of auditory displays.

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. This results in a deeper immersion into the virtual environment and allows a better perception of the conveyed information. This is a necessity for being able to correctly decode the information 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. 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. A more advanced authoring tool is currently under development which allows both, a visual and an auditory-only interface to design these 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.

6. REFERENCES

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

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

[3] John Alton. *Painting With Light*. University of California Press, 1995.

[4] Durand R. Begault. *3D Sound - For Virtual Reality and Multimedia*. AP Professional, 1994.

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

[6] A. S. Bregman. *Auditory Scene Analysis: The perceptual Organization of Sound*. MIT Press, 1990.

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

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

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

[10] C. Frauenberger and M. Noisternig. 3d Audio Interfaces for the Blind. In *ICAD*, 2003.

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

[11] 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.

[12] 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.

[13] John Garas. *Adaptive 3D Sound Systems*. Kluwer Academics, 2000.

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

[15] Pin Interactive. Terraformers. Published by Pin Interactive, 2003. PC.

[16] Christopher Lewis. Audio Formula 1. Published by Christopher Lewis, 2003. PC.

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

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

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

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

[21] Modest Mussorgski. Pictures at an Exhibition, 1874.

[22] 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.

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

[24] Sergej Prokofjew. Peter and the Wolf, 1936. Op 67.

[25] Maurice Revel. Pictures at an Exhibition, 1922.

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

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

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

[29] 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.

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

[31] Isao Tomita. Pictures at an Exhibition, 1975.

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

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

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