1 INTRODUCTION

Imagination is more important than knowledge. — Albert Einstein

WE live in a sensual environment; a world that is filled and stimulated through objects that can be seen, heard, touched, smelled and some of them even tasted. Out of these five senses, vision is our strongest and most prominent, but we often fail to appreciate how good we really are at interpreting sounds and noises. Sound is everywhere around us and perceived through the sense of hearing. It enriches our visual environment and accompanies us through our daily routines, informs us about current events and occurrences, and assists us in performing tasks and duties by providing auditory feedbacks. Behind the ability to derive abstract information from auditory signals lies another, more emotional layer. In the form of music and auditory reminisces, it affects and touches us in a very deep and emotional way, and has therefore – in certain cases – a much stronger influence, superior to that of vision.

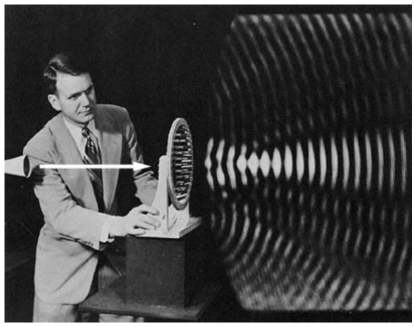
Sound and acoustics have both been practiced in the form of music and singing throughout human history, but had not been studied in greater detail until Pythagoras and Aristotle. Both looked at sound and acoustics from the viewpoint of science and discovered the very basic fundamentals of harmonics and the propagation of sound waves. Several of these effects, including the interactions (reflections) of sound waves with different materials and objects were already known at this time to Greek and Roman architects, who designed their theaters and halls with *good acoustics* in mind; the early beginnings of architectural acoustics. Today's physical definition of sound describes it as a disturbance of mechanical energy that is propagating through matter in longitudinal waves. The foundations for this understanding were laid during the time of Scientific Revolution by Galilei, Mersenne and Newton, who studied not only the physics of sound wave propagation, but also its psychological effects in perception. Newton later derived the relationship for wave velocity in solid objects, and thereby marked the beginning of the physical understanding of acoustics as we still have it today (Newton, 1687). In the 19th century, acoustics and sound propagation were studied by Helmholtz in Germany and Lord Rayleigh in England, who later also compiled the first monograph in this still very young field of research: *The Theory of Sound* (3rd Baron Rayleigh, 1877/1878).

In today's technological and information-driven environment, sound and acoustics have many applications, yet some areas are still unexplored. To shed some light – or sound waves – on some of these areas is the scope of this thesis. Sound and auditory perception have some very unique features, different to seeing and the perception of light, which shall be explored, examined and exploited for new applications and new possibilities. As characterized by the quote at the beginning of this chapter, imagining the unexplored possibilities is always the first step in the beginning of new research. In this thesis, it starts right in the next section, focussing on an INTERACTION WITH SOUND.

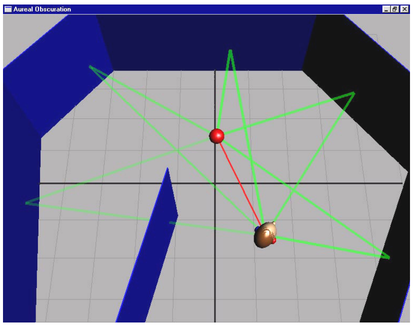
1.1 INTERACTION WITH SOUND

Envisioning an INTERACTION WITH SOUND exhibits an interactive, user-centered dialog that is located within an auditory environment. The conveyance of information in a so called *Auditory Display* focusses primarily on sound and acoustics, that means, abstract information is represented and *displayed* using auditory means and sound signals (Kramer, 1994). Although this seems difficult and applicable to a few cases only, these

(a) Léon Theremin is playing the *Theremin*, one of the first electronic music instruments. (Theremin, 1924)



(b) An acoustic lens visualizing sound waves and acoustic amplification patterns. (Knight, 1960)



(c) Wavetracing: A visualization of *Aureal's* realtime ray tracing system for room acoustic simulations. (Ramelet, 2000a)

(d) The ReacTABLE*: A tabletop designed musical instrument (synthesizer). (Kaltenbrunner et al., 2006)

Figure 1: Several examples for the *Interaction with Sound*.

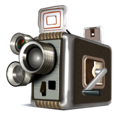
techniques are employed within a variety of applications. Most people are unaware of how important auditory information can be, and how well humans really are in the interpretation of acoustics and sound signals:

> *"Sound plays an integral role in our everyday encounters with the world, a role that is complementary to that of vision."* (Gaver, 1989)

An short example for such an auditory display is the acoustic notification for a new received email. A similar *Auditory Icon* (sound) is employed in all major mailing applications, and if we hear this sound, we immediately know that new email has arrived. However, an INTERACTION WITH SOUND describes a dialog with a broader application that can be employed for an acoustic display of various information and tasks.

The use of sound for an abstract information representation has several advantages. Most notably the possibilities for a subconscious presentation and reception, in which the information is presented in a way that the user can perform other – possibly visual – tasks with full performance, but is notified acoustically about occurring events. Existing graphics-based visualization systems can be enhanced by an auditory display, in a way that a multi-modal presentation is employed to increase the understanding of the data and the performance of the user. Although the required sound signal processing is in certain cases relatively complex, the overall hardware requirements are less demanding. This makes an application of an INTERACTION WITH SOUND very suitable for portable hardware and location-aware tasks. The design of the interactive components and the way an interaction is performed with these auditory environments is required to be audio-centered. This means that certain listening and interaction techniques from the real world should to be mimicked and integrated into the system, while focussing on an audio-only display of this information. Here the application of speech recognition and synthesis plays a large role, as speech permits for a very abstract, but also very direct presentation of information.

The design of an interactive auditory dialog system requires a semantic/semiotic component for the content, and the actual part of interaction. The fundamentals for an INTERACTION WITH SOUND are found in the area of *Auditory Displays*, but also in the visually dominant field of *Visualization* (Kramer, 1994). Schumann and Müller define three basic principles that characterize a *good* visualization by describing the expressiveness, efficiency/effectiveness and adequacy of the display (Schumann and Müller, 2000). These principles can directly be transferred to an auditory presentation of information and the design of auditory displays as well. To find and identify these similarities and shared aspects with other areas is one goal of this research. The results are used to define rules and guidelines for the design, authoring and application of 3D VIRTUAL/AUGMENTED AUDITORY ENVIRONMENTS.



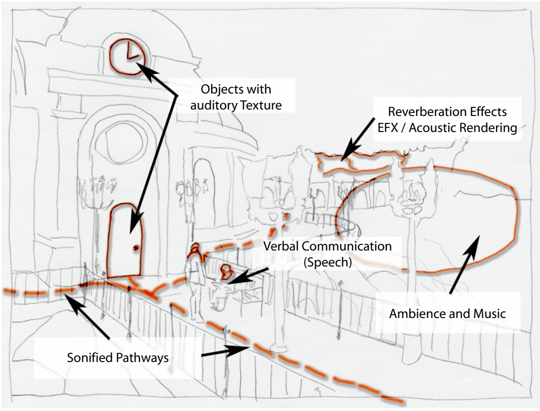
Playing the Theremin.

Examples of an INTERACTION WITH SOUND, fundamental to this research, can be seen in Figure 1. The first example in Figure 1a shows the Russian inventor Léon Theremin playing an electronic instrument that he created (Theremin, 1924). It is played by using hand gestures, with two metallic antennas sensing the hands position to control an oscillator and a volume effect. The resulting sound is very unique and the instrument is currently experiencing a renaissance. The second example in Figure 1b visualizes sound waves using an acoustic lens: *"This new technique of studying sound demonstrates the focusing effect of an acoustical lens on sound waves issuing from the horn at extreme left."* (Knight, 1960). The lens was developed at the *Bell Telephone Laboratories*, who experimented at this time with technologies for creating personal listening spaces. The study of human auditory perception and sound wave propagation accounts for a broad discussion in this research due to its high significance. Figure 1c displays the principle of *Aureal's* revolutionary wavetracing technology (Schneider and Muschett, 1998; Aureal, 2000; Ramelet, 2000a). In 1998/1999 Aureal developed the first PC sound hardware capable of performing simple, yet very realistic room acoustic simulations. This hardware employed a realtime technique called *Wavetracing* that used actual geometry to perform a more realistic – physically-sound – acoustic simulation. The ReacTABLE* system, as shown in Figure 1d, is an intuitive virtual instrument that simulates an analog synthesizer (Kaltenbrunner et al., 2006). It is played by hand using a tabletop design approach and employs computer vision techniques to interact with the instrument. Although these four examples seem very diverse, they all represent different areas of this research that are combined in an *Interaction with Sound*. Several of these ideas are referred to in later chapters, in which they are applied for an interaction and exploration of 3D virtual auditory environments.

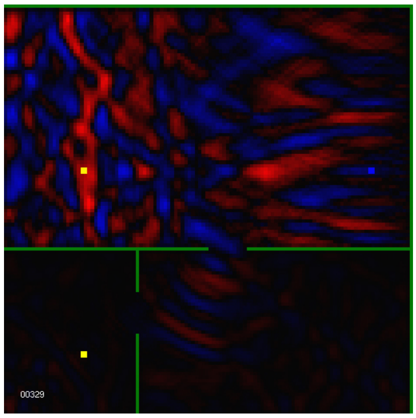
A Demonstration of Aureal's Wave-tracing Technology.

A Demonstration of the ReacTABLE System.*

The above examples show that sound and an auditory display of information has many advantages, but also several drawbacks. Some of the larger issues – compared to vision and graphics – are a limited signal bandwidth and a serial perception of information. Sound is only audible at the particular moment in time in which it is created. The ease with which humans are able to derive information from figures and drawings, but also how we perceive and listen to our environment, was learned and trained over centuries. This knowledge is very important, not only for the design of auditory user interfaces, but also for the encoding of information and its presentation. Although the visual senses have always dominated the perception of information, a proper training of the auditory skills permits an increase in performance to a level that is equal to the visual system (Gaver,



(a) A Definition of 3D virtual auditory Environments that visualizes the Concept and the Requirements to design and sonify 3D auditory Spaces, as well as denotes Techniques for an adequate Interaction.



(b) An Animation time frame that shows an efficient GPU-based Implementation of a physically correct Sound Propagation Model based on differential Equations.

Figure 2: Two Contributions of this Thesis.

1989). In several experiments this observation could be confirmed with the evaluation of own prototypes and example implementations.

Focussing on the advantages and possibilities of auditory display systems, many interesting and *cool* applications emerge capable to provide solutions for a variety of problems. The willingness to explore new, and sometimes unconventional user interfaces, is strongest among the communities of entertainment and edutainment software users (Yatim and Masuch, 2007). Especially computer games often experiment with new interaction paradigms, which later – if successful – are often applied to other, more conservative, software applications. The possibilities for using sound are not restricted to the sonification of abstract scientific data sets alone, but can, with the areas of enter- and edutainment in mind, be applied to highly immersive interactive auditory games and learning systems. These techniques can also be used to design augmented audio reality applications for the task of guiding tourists, or to aid the visually impaired in finding pathways and for the performance of daily routines. The same techniques can furthermore be used in ubiquitous and mobile computing scenarios to enhance our local environment by an *ambient intelligence*. The possibilities are not limited, but in order to realize these applications, more research in the directions of auditory perception and information presentation, especially for 3D virtual (augmented) environments is required. This shall be the scope of this research.

1.2 OBJECTIVES AND CONTRIBUTIONS

The preceding section discussed several interesting questions regarding the research for an *Interaction with Sound*. Current and existing research has so far not studied 3D virtual auditory environments in greater detail, especially not from a unifying perspective and with an audio-centered design in mind. The majority of related work employs auditory environments and 3D auditory displays to enhance the depiction of visual information, whereas this research explicitly focusses on audio-only representations to exploit the benefits of such a display. One of the objectives of this research is to explore possibilities, and to answer questions around the design and application of 3D auditory display systems and for 3D virtual auditory environments. The research thereby provides an extensive overview of the area and performs detailed studies on selected applications. A goal is to provide metrics and software construction kits, using which a user would be able to select appropriate techniques of information sonification and interaction that are both consistent with the task of application. The main objectives can be summarized as:

- An analysis of audio used in entertainment computing and the study of 3D virtual (augmented) auditory environments,
- The design of suitable sonification and interaction techniques for an exploration of and interaction with 3D auditory environments,
- A study of authoring and design guidelines for the construction of 3D virtual (augmented) auditory environments, and
- An exploration and evaluation of several example scenarios.

During the years of research, all of these areas have been studied in detail and many example applications were developed and evaluated. The main focus was thereby centered around the definition, analysis and classification of 3D virtual auditory environments along with several related application scenarios. Contributions have been made on several – partially very diverse – areas and fields of research:

- A refinement of existing 2D/3D data, image and volume sonification techniques.
- A finer definition of 3D virtual auditory environments using a *non-realistic* auditory scene design, as well as advancements regarding the sonification of 3D scene information and the development of 3D spatial interaction techniques.
- An advancement of augmented audio reality in terms of 3D spatial interactions and user-orientation/positioning techniques, as well as an evaluation of example scenarios using an efficient, self-developed low-cost system.
- A derivation of 3D scene authoring and design guidelines which are integrated into a 3D auditory scene authoring environment.
- A development of graphics-based 3D scene auralization and sound rendering techniques that improve existing solutions in terms of quality and efficiency.

Figure 2 shows exemplarily two contributions of this research. Figure 2a visualizes a description of a 3D virtual auditory environment in which objects of importance are highlighted. These objects have been assigned with interactive auditory textures, which allow an intuitive exploration and sonification of this 3D auditory scene. Figure 2b displays another, more technical contribution of this thesis, and shows an animation frame of an interactive wave-based sound propagation system that exploits computer graphics hardware for an efficient and more accurate sound simulation. Although the two results seem to have – at a first glance – not much in common, yet both rely on each others principles. As the interaction with audio-only environments is difficult and requires accurate techniques for sound rendering and simulation that currently available APIs can only partially fulfill, additional research was required to design an efficient, yet highly realistic, sound rendering and simulation system.

Graphics-based Sound Simulations.

1.3 PERSPECTIVES

From the perspectives described in this section, several directions of research are opening up. The one chosen in this thesis will continue towards 3D auditory displays and the design of 3D virtual/augmented auditory environments. Along this way not only the theoretical foundations are explained and discussed in detail, but also several fundamental techniques and prototypic applications are developed and explored. Chapter 2 starts with a theoretical and at the same time very global perspective of an INTERACTION WITH SOUND, and classifies the thesis' topic within its larger areas of research in computer science. Chapter 2 also presents the major scientific goals of this research and discusses issues of methodology and special requirements for an effective implementation. At the end of Chapter 2, an overview of the thesis' organization is provided, which describes the content and the contribution of each chapter.