# 2 ANALYSIS

*The way of the Samurai is one of immediacy, and it is best to dash in headlong. — Tashiro Tsuramoto*

THE goal of this research is to explore and foster 3D virtual/augmented auditory environments – especially within the domains of entertainment and edutainment applications – by using tools and techniques from computer games, 3D auditory displays and scientific data sonification. This chapter is used as a starting point for this research and introduces the most important concepts and terminologies, as well as classifies the thesis topic and analyzes the major scientific goals. Furthermore, this chapter provides an overview of the methodologies employed and discusses some of the requirements and settings necessary to obtain the identified objectives. As the research employs techniques from several different areas – not all from within computer science – this chapter aims at providing a broader perspective in which to present the topic as a whole and to unify the domains involved. In reference to the quote at the beginning of the chapter, it is often favorable to not ponder too long with a task at hand, but to *dash in headlong*.

## 2.1 ANALYSIS AND CLASSIFICATION

The topic of *Interaction with Sound* overlaps with several domains in computer science and its associated fields. Although the research is centered around sound and acoustics, methods of Human-Computer Interaction (HCI) play a major role in evaluating the quality and usability of the designed systems and applications. The heart of this thesis are 3D VIRTUAL AUDITORY ENVIRONMENTS, around which all research and discussions are arranged. With the means of perceiving information consciously reduced to hearing alone, the interaction and sonification techniques that are required have to be designed to compensate for some of the missing visual cues. These differences in perception, although resulting from missing pieces of information, yield to a new way of information presentation that offers alternative ways to describe 3D virtual environments. Depending on the task and area of application, the requirements therefore vary and are expressed through the different sonification/interaction techniques and the alternative auditory presentations used. Some of the basic techniques can also be applied and evaluated using the sonification of 2D and 3D scientific data sets. The focus, however, lies on 3D auditory spaces and a combination with concepts and techniques from 3D computer games to define interactive 3D virtual/augmented auditory environments.

Based on this short analysis, the fundamental areas of this research can be identified as the following five domains:

- Human-Computer Interaction (HCI),
- Auditory Displays,
- Audio and Acoustics,
- 3D Virtual and Augmented Environments, and
- Entertainment and Edutainment Applications (eg. Computer Games).

Each of these topics represents an individual research domain or area of application, but combined, they shape the foundation of this research with 3D virtual auditory environments at its center. Overlapping areas have to be identified and analyzed in order to successfully combine all topics and unify their strengths and advantages. So far, a clear definition of 3D virtual auditory environments that integrates both virtual reality (VR) and 3D auditory display systems is still not available, especially not from the perspectives of entertainment and edutainment applications. Although the term is used and known, it is employed in an ambiguous manner and defined from varying positions. Therefore, a clear definition with respect to the above mentioned research domains is required. This definition must also include 3D scene sonification and interaction techniques to convey information of the virtual environment to the user, and to input user information into the system. The goal of this thesis is to develop and establish 3D auditory environments as an *equal* to visual environments. In the following, each of these five research domains is shortly introduced and discussed in their own context, as well as in respect to their possible contributions and potential for 3D virtual auditory environments.

HUMAN-COMPUTER-INTERACTION defines the upper layer of abstraction of this research. It is the interdisciplinary study of interfaces and the interaction between humans and machines, and is concerned with their design, implementation and evaluation. With the focus on an auditory perception and communication, it is used to study the acoustic conveyance and presentation of data and information. In addition, HCI is also responsible for the design and development of efficient and effective techniques for an intuitive interaction with 3D auditory spaces. The goal will be the definition of building sets and guidelines, which allow a task-dependent selection of suitable sonification and interaction techniques. Using software evaluation techniques, these selections, as well as the prototypes developed, can later be evaluated and their functionality be assessed. Further contributions are concerned with the development of authoring and design guidelines to provide an adequate balance between an applications function and its aesthetics in presentation.

AUDITORY DISPLAYS represent the closest area of related work and at the same time provide the basic methods and technologies necessary to design and implement the ideas of this research. Auditory displays are similar to visual displays in the respect of *displaying* information, except that auditory displays are based on auditory means and primitives to convey abstract information and data. Many of the thesis' concepts and ideas are directly related to existing approaches emanating from this area. The approaches developed within this thesis can therefore be very well compared with related and existing concepts and techniques.

AUDIO AND ACOUSTICS are used in this research as the primary means of communication, and are therefore the focal point of all techniques, methods and applications developed. As the acoustic perception of information is quite different to visual seeing, a large portion of this research concentrates on an efficient presentation using a non-realistic auditory design, ie. an auditory representation that deliberately departs from a physically correct display and concentrates on a perceptual presentation. A direct requirement for this task are efficient techniques for for a high-quality acoustic rendering and 3D sound synthesis.

3D VIRTUAL AND AUGMENTED ENVIRONMENTS contribute both the platform and the stage to this research. Virtual reality and 3D virtual/augmented environments are very common in computer graphics, and therefore many of the visual display and interaction designs available can be transferred into their auditory counterparts. The challenge of this research is to map these concepts towards auditory presentations and onto auditory primitives respectively. This requires rules and guidelines for the mapping of abstract information, as well as suitable 3D spatial interaction techniques for the design and authoring of 3D virtual auditory environments.

**ENTERTAINMENT AND EDUTAINMENT APPLICATIONS** serve as basis and test ground for the majority of examples and implemented prototypes. Advantages of using computer games are the great number of possible scenarios and a high suitability to evaluate new and unconventional interface designs and interaction techniques. Furthermore, computer games and edutainment applications can be used to assess the quality of additional attributes, such as storytelling, narration and immersion. In a future setting, the here developed concepts and techniques for an interaction with auditory environments should be applicable to scenarios beyond entertainment and edutainment as well.

This domain analysis and the classifications of the topic allow now a formulation of the most important research questions and scientific objectives. With the focus on 3D virtual auditory environments and the goal to provide techniques for an intuitive 3D scene sonification and interaction, the two primary research objectives are:

- The analysis, survey and classification of 3D virtual auditory environments in terms of presentation, realism, interaction and area of application, as well as
- The development of metrics and techniques to define and select suitable and task-dependent 3D sonification and interaction techniques for the tasks of 3D scene exploration, the conveyance of abstract object/scene information, as well as to perform 3D spatial object/scene interactions.

A major challenge is the detailed analysis and comparison of visual and auditory environments regarding the conveyance and representation of information, as well as the development of methods for an audio-centered selection, object emphasis and scene/object interaction. As both environments represent and display abstract information using either visual and/or auditory means, several concepts from information visualization might be applicable and transferable to the auditory realm as well. Problems will not only arise due to the differences in perception, but also due to the different environments that both senses describe. With this in mind, the secondary objectives for this research are identified as:

- The analysis of audio in entertainment computing and in auditory displays.
- A detailed comparison of visual and auditory 3D environments with a focus on presentation and interaction.
- The proposition of techniques and guidelines for the design of enhanced non-realistic 3D virtual auditory environments.
- The design of 3D scene sonification and spatial interaction techniques, suitable for an exploration and interaction within 3D virtual auditory environments.
- The conception of authoring techniques and guidelines for the design of 3D virtual auditory environments.
- An evaluation regarding an applicability and implementation of 3D augmented audio reality.
- The design and implementation of an interactive audio framework, applicable to several different example scenarios and areas.

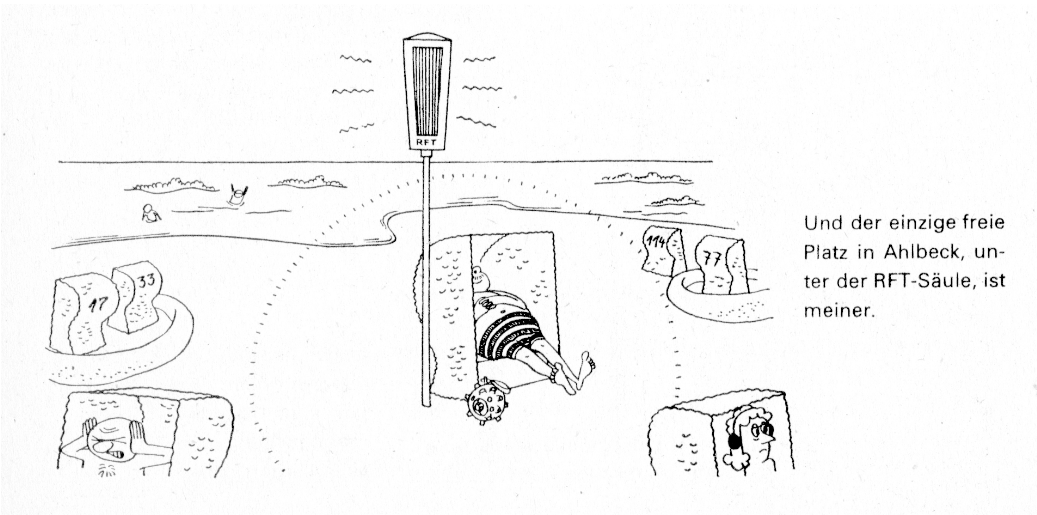


Figure 3: The auditory hedgehog: *"And the only free place in Ahlbeck, under the RFT speaker, was mine."* (Schmitt, 1968).

## 2.2 OBSERVATIONS

A major difference between an auditory and a visual perception of information is that auditory information is constantly audible during the actual display. Unlike eyes, which can be closed, every acoustic signal in the local environment that is audible *will be* audible. This problem makes a close proximity installation of several auditory displays difficult if the information is presented over a speaker array and not through headphones. Therefore, the research in this thesis employs headphones as actual *physical* auditory display in almost all examples.

The ever growing flow of noise and auditory output in our modern western civilization serves as an unpleasant experience to more and more people. Schmitt, a cartoonist born in the GDR, devised here several interesting ideas in his book *"Cartoons in the service of science"*, which are not only applicable to sound and acoustics (Schmitt, 1968). His auditory hedgehog *bends* sound waves into straight lines and is thereby creating a silent spot for his own personal enjoyment, compare with Figure 3. Using modern noise canceling headphones[1] and efficient acoustic insulations, this experience can be partially recreated.

Aside this interesting direction of research, the following two sections focus more on the objectives that were devised at the beginning of this chapter. A preliminary concept for the research along with necessary procedures is formulated and later extended by an initial requirements analysis that is concerned with an implementation of the above objectives.

### 2.2.1 *Procedures*

With the analysis completed and the scientific goals identified, a first concept can be drawn that summarizes the research and a realization of the envisioned objectives. But first, the major research questions that are to be answered by this thesis are defined in the form of hypotheses. The underlying theories will be examined in greater detail by

1 http://www.noisefreeheadphones.com/

the individual chapters and are later evaluated and either confirmed or rejected at the final discussions in Chapter 10.

The initial questions for the development of a first concept and approach towards 3D virtual auditory environments are:

- What can be described with 3D virtual auditory environments, and what are the possible applications?
- What level of realism is required and achievable?
- How important is 3D sound spatialization, and what are the individual perceivable differences?
- What is the most intuitive way to acoustically display information, and how can one interact with such an auditory environment?
- How can a real-world environment be combined with an artificial auditory space?
- What are the requirements for the design of 3D auditory environments, and where are the limitations?
- How accurate is the human perception of 3D sound and environmental acoustics, and what are the requirements for a realistic acoustic rendering and sound simulation?
- How suitable are 3D auditory environments for gameplay, and what level of immersion can be achieved?

Several steps are required to answer these questions. The first one is an extensive analysis of existing and related work along with a detailed study of auditory perception and acoustical effects. These results will allow a first approximation in the direction of what can be achieved and how 3D virtual auditory environments have to be designed and constructed for an effective and expressive auditory display of information.

The second step will involve a closer observation of 3D auditory environments in order to derive suitable task-related sonification and interaction techniques. A possible approach is here to employ a collection of generally suitable techniques and to use metrics to decide upon and select appropriate methods of interaction that are consistent with the sonification goal specified. Potential difficulties may arise from auditory perception, insufficient sound rendering and synthesis techniques, as well as from problems of mapping abstract information onto acoustic primitives for an auditory display. Depending on their severity, more research has to be conducted in either of these areas to improve the perception, sonification and interaction with 3D auditory environments.

In a third step, the idea of combining a virtual auditory environment with real surroundings moves into focus. The question is here whether or not it is possible to project a convincing artificial auditory environment onto an existing real one: enhancing and augmenting a natural auditory environment through artificial sound signals. Several ideas of promising applications are emerging and range from guiding systems for the visually impaired to interactive systems for an auditory storytelling. As the direction of research loosely points towards entertainment and edutainment applications, questions regarding an auditory gameplay and the achievable degree of immersion arise. Due to the missing visual cues, which are now substituted by the players own imagination, it can be assumed that an audio-only presentation achieves a higher level of immersion. This would make auditory environments very applicable for narration and adventure-based computer games.

Besides the development of prototypes and applications, it is also important to know the limitations of 3D auditory environments and auditory displays. Therefore, an additional emphasis lies on devising rules and guidelines for an authoring and design of interactive 3D virtual/augmented auditory environments.

### 2.2.2 *Requirements*

In order to attain the above objectives and to answer the research questions described, several requirements have to be fulfilled. The initial comparison of visual and auditory presentations includes a review of related work and similar approaches, but also the conception of small prototypes for a more detailed study of auditory perception. These prototypes are integrated into a more sophisticated audio framework, which will be utilized throughout the research for the evaluation of the proposed concepts and ideas. This framework might also be used for creating prototypic applications, but its main task is to evaluate specific sonification and interaction methods for the design of a task-dependent construction kit. Such a framework can be designed analogously to the layout of computer game engines, which support a visual and auditory display of information, as well as an interaction with 3D virtual environments. With the focus on 3D virtual auditory environments, a first estimate of the requirements is:

- A 3D polygon-based scene management system which supports collision detection,
- Highly efficient and accurate 3D sound rendering and room acoustic simulation techniques, as well as
- Several possibilities to implement 3D interaction paradigms, such as exploration, orientation and navigation, and 3D object/scene interactions.

An initial design approach that can be employed for the first evaluations and test scenarios can be developed using standard APIs and libraries, such as OpenGL, OpenSG, OpenAL and DirectX. This prototype can later be extended to support more advanced 3D spatial interaction techniques using 3D user tracking hardware. Another interesting possibility emerges with the replacement of the standard sound rendering API (eg. OpenAL), which can be substituted through a dedicated system that supports a higher accuracy and efficiency. This discussion of a requirements analysis is continued in later chapters, which explicitly concentrate on system design and implementation.

All of the research questions and objectives discussed are further examined and studied in their respective chapters. The next section provides an overview of the structure of the thesis, as well as a summary of each chapter.

## 2.3 THESIS OUTLINE

The thesis is organized and structured into 10 chapters, of which the first chapters are of more introductory nature and lay a foundation for the coming design of 3D virtual auditory environments. Code examples and implementation details can be found in Appendix A, but are also integrated within the algorithmic discussions of the respective chapters. Specific examples and results are discussed throughout the thesis, but are additionally summarized in Chapter 9 and Appendix C. Chapter 9 provides additional details and examines several prototypic implementations and case studies from a chapter-overlapping perspective.

The thesis is organized as follows:

FUNDAMENTALS — Chapter 3 and Chapter 4 illustrate and discuss the fundamentals of this thesis, as well as examine related work and similar systems. Chapter 3 concentrates on the fundamentals by introducing important concepts, terminologies and applications, along with the necessary details of auditory perception and sound signal processing. The goal is to familiarize the reader not only with physiological and psychological, but also with important technical aspects of sound perception and rendering. The chapter concludes with a discussion of audio in entertainment computing and examines several existing applications.

AUDITORY DISPLAY — After the discussion of basic fundamentals, Chapter 4 concentrates on the application's side and analyzes 2D and 3D auditory display systems and sonification techniques. The chapter examines existing concepts and implementations, with a focus on an intuitive and efficient sonification of information. The chapter also discusses issues and open problems of an auditory user interface (AUI) design, as well as provides a general overview on the subject of auditory presentation and display using several examples.

AUDITORY ENVIRONMENTS — The following Chapter 5 through Chapter 8 contain the main contributions of this research. Chapter 5 starts here with a discussion of 3D virtual auditory environments. After an abstract definition of virtual reality and 3D auditory environments, the chapter concentrates on an intuitive auditory scene design and develops the concept of a non-realistic auditory scene presentation. Using this approach, the remaining sections discuss and implement techniques for 2D/3D data and 3D scene sonification, as well as develop intuitive methods for 3D spatial interaction. The chapter concludes with a discussion and the design of an audio framework that allows an implementation and evaluation of the described concepts.

AUGMENTED AUDIO — Chapter 6 directly connects to Chapter 5 and extends the framework developed towards an augmented and mixed reality design. In this chapter, a low-cost augmented audio reality system is described and developed, and applications that combine a real-world environment with an artificial auditory scene are discussed. Although augmented audio reality possesses many advantages and possibilities, it also exhibits several difficulties that are addressed within this chapter. This includes an extension of the previously devised sonification and interaction techniques, as well as the development of efficient methods for user tracking and positioning. In alliance with Chapter 5, the last section extends the existing audio framework and discusses possible applications for augmented audio reality.

AUTHORING AND DESIGN — While the two previous Chapter 5 and Chapter 6 were more concerned with an implementation of 3D virtual/augmented auditory environments, Chapter 7 provides a closer look regarding issues of authoring and design. As the auditory channel has a small bandwidth and information can only be perceived in a serial manner, these issues are of high importance. The chapter therefore concentrates on the development of rules and guidelines for the various authoring tasks, and also devises an authoring environment that implements these concepts exemplarily.

ACOUSTIC RENDERING — Chapter 8 is motivated through the special requirements for the design of auditory environments as discussed in Chapter 5 and Chapter 6. Chapter 8 discusses and implements efficient and accurate techniques for an acoustic rendering of 3D spatial auditory environments. After a short introduction of the subject, several techniques for the spatialization of monaural sounds, as well as for the simulation of room acoustics are explained. Due to several qualitative and quantitative requirements, an efficient implementation of these techniques using computer graphics and graphics hardware is provided, as well as compared with state-of-the-art implementations. The chapter concludes with the discussion of a sound engine that not only exploits computer graphics hardware, but also transfers graphics-based designs towards 3D acoustics and 3D sound simulations.

CASE STUDIES — Several examples and results are discussed throughout their respective thesis chapters. Chapter 9 summarizes the majority of these implementations and presents applications and case studies from a broader perspective. The majority of applications and examples that are discussed within this chapter were examined and evaluated through user studies. Their results are discussed individually for each area of application. The chapter focusses especially on the evaluation of sonification and interaction techniques for 3D auditory scenes, but also discusses 2D/3D data sonification, audio-only computer games, augmented audio reality applications and interactive audiobooks. The chapter concludes with an additional analysis of the sound rendering techniques developed in Chapter 8, and summarizes the results from all applications.

CONCLUDING REMARKS — Finally, the thesis is summarized and conclusions are drawn in Chapter 10. The initial goals are compared with the results achieved, and propositions are developed to describe the essence of this research. The chapter also discusses ideas and possibilities for future improvements that could not be addressed in this thesis.

APPENDICES — Succeeding the chapters of this thesis are three appendices, which provide additional code examples (Appendix A), discuss the user studies and evaluations in greater detail (Appendix B), as well as list and describe the examples contained on the accompanying DVD (Appendix C).