UNIVERSITY OF OSLO
Department of Informatics

Follow the Sound

Design of mobile spatial audio applications for pedestrian navigation

Master thesis

Joakim Bording

1. August 2012

Abstract

Auditory displays are slower than graphical user interfaces. We believe spatial audio can change that. Human perception can localize the position of sound sources due to psychoacoustical cues. Spatial audio reproduces these cues to produce virtual sound source position by headphones. The spatial attribute of sound can be used to produce richer and more effective auditory displays.

In this work, there is proposed a set of interaction design guidelines for the use of spatial audio displays in a mobile context. These guidelines are inferred from psychoacoustical theory, design theory and experience with prototype development. The horizontal front arc is presented as the optimum area for sound localization, and the use of head- or body-tracking is stated to be highly beneficial.

Blind and visually impaired pedestrians may use auditory displays on mobile devices as navigation aids. Such aids have the potential to give visually impaired access to the environment and independence of movement. Custom made hardware is not always needed, as today's smartphones offer a powerful platform for specialized applications.

The Sound Guide prototype application was developed for the Apple iPhone and offered route guidance through the spatial position of audio icons. Real-time directional guidance was achieved through the use of GPS, compass sensor and gyroscope sensor. Spatial audio was accomplished through the use of prefiltered audio tracks that represented a 360° horizontal circle around the user. The source code of this prototype is made available to the community.

Field tests of the prototype were done with three participants and one pilot tester that were visually impaired. One route was navigated with the help of the prototype. Interviews were done to get background information on navigation for visually impaired pedestrians. This was done to see how the prototype was received by visually impaired test users and what can be done to improve the concept in later development.

Even though the prototype suffered from technical instabilities during the field tests, the general responses were positive. The blind participants saw potential in this technology and how it could be used in providing directional information. A range of improvements on the concept has been proposed.

Acknowledgements

I would like to take this opportunity to show my gratitude to those people that have helped me in the work with this thesis.

My supervisor **Jo Herstad**, at the Department of Informatics at the University of Oslo, was a highly valuable advisor. He provided enthusiasm, insight and guidance when I needed it most. This thesis will not have been the same without him.

Jo got me in touch with **Magne Lunde** and **Morten Tollefsen** at Media Lt. They deserve credit, for inspiring me to pursue the task of pedestrian navigation in this thesis. They shared their knowledge in the needs of people that are visually impaired and helped in finding participants for the study conducted in this work.

I would like to thank the three participants, and the two pilot testers that used their valuable time to test my prototype and share their experiences with me. Your contributions were of great importance to me.

Thanks to **Paul Bennun** from Somethin'else for taking the time to answer questions about the inner workings of the Papa Sangre game. Thanks to **Ellen Sundh** from Society46 for her advice on how to implement a key feature in the development of the prototype. Also, thanks to the fine people at the reading hall in the 7th floor at Ole Johan Dahls house, especially **Emanuele Lapponi** and **Mads Jahren** that made long days feel shorter.

Many people have helped me in the writing of this thesis. Thanks to **Jan Kåre Bording**, for his precise critiques; **Jan Bording**, for his broad insight and bright ideas; and **Bente Bakke**, for providing vital feedback and showing great patience, even though her boyfriend spoiled most of their vacation on thesis writing.

My years at the university have helped me to understand the world around me and myself. I would like to thank my family for having inspired me to pursue higher education and the knowledge it entails.

Joakim Bording
Oslo, 30. July 2012

Contents

1 Introduction **1**
- 1.1 Research Area 3
- 1.2 Motivation 3
- 1.3 Research Questions 4
- 1.4 Problem Analysis 5
- 1.5 Contributions 6
- 1.6 Overview 7

2 Theory **9**
- 2.1 Human Computer Interaction 9
  - 2.1.1 Interaction Design 9
  - 2.1.2 Frameworks in Design 12
  - 2.1.3 Auditory Displays 14
- 2.2 Spatial Audio 16
  - 2.2.1 Psychoacoustical Cues in Sound Localization 17
  - 2.2.2 The Cocktail Party Effect 20
  - 2.2.3 Stereo Audio 20
  - 2.2.4 Binaural Recordings 21
  - 2.2.5 Head Related Transfer Functions 21
- 2.3 Mobile Hardware and Software Support 22
  - 2.3.1 Hardware Support 23
  - 2.3.2 Software Support 24
  - 2.3.3 Commercial Devices with 3D Audio Support 25
- 2.4 Projects Utilizing Spatial Audio 26
  - 2.4.1 Commercial Applications 26
  - 2.4.2 The UCSB Personal Guidance System 29
- 2.5 Visually Impaired Pedestrians and Navigation 32
  - 2.5.1 Orientation and Mobility 33
  - 2.5.2 Mobility Aids 34
  - 2.5.3 Orientation Aids 35
  - 2.5.4 Guidelines for Designing O&M Aids 36

3 The Sound Guide Case **39**
- 3.1 Concept and Vision 39
- 3.2 Software Developments 40
  - 3.2.1 Technologies Used 40
  - 3.2.2 Implementation 42

3.2.3 User Interface Design 44
3.3 Comparisons with other Applications 48
3.4 Use Scenario 49

4 Methods **51**
4.1 What is a Method 51
4.1.1 Quantitative Methods 51
4.1.2 Qualitative Methods 52
4.1.3 Terms and Constructs 52
4.2 Methods 53
4.2.1 Semi-structured Interviews 53
4.2.2 Field Studies and Observation 54
4.2.3 Content Analysis 54
4.3 Ethical Considerations 55
4.4 Research Design 55
4.4.1 Participants 55
4.4.2 Time, Location and Logging 56
4.4.3 Test of Sound Localization Accuracy 56
4.4.4 Interview 57
4.4.5 Field Test 57

5 Results **59**
5.1 Test of Sound Localization Accuracy 59
5.2 Interviews on Navigation 60
5.2.1 Information in Navigation 61
5.2.2 Navigation, Orientation and Mobility 62
5.2.3 Methods of Navigation for Visual Impaired 62
5.2.4 When Navigation Fail 65
5.2.5 Use of Cellphones 65
5.2.6 Use of Headphones 66
5.3 Usability Tests of the Sound Guide 66
5.3.1 Technical Performance 66
5.3.2 Interpretations of Audio Display 67
5.3.3 Interaction 69
5.3.4 External Influence 70
5.3.5 User Experiences 71

6 Discussion **75**
6.1 Research question 1 75
6.1.1 Spatial Layout 75
6.1.2 Sound Design 78
6.1.3 Interaction 81
6.1.4 Sounds in the Environment 81
6.1.5 Realism or Enhancements 82
6.1.6 Individuality in Perception 83
6.1.7 Implementation 85
6.1.8 Summary 87
6.2 Research question 2 88

6.2.1 The Sound Guide prototype 88
6.2.2 Directions by Compass 88
6.2.3 Sound Design 90
6.2.4 Audio Filtering 96
6.2.5 Implementation 98
6.2.6 Independence of Movement 101
6.2.7 Summary 103

7 Conclusion **105**
7.1 Final Results 105
7.2 Future Work 106

List of Figures

1.1 A broad range of theory is used in this case in the attempt to give a satisfactory foundation to answer the research questions stated. 8

2.1 This illustration show how occlusions by the head can lead to intensity differences between sounds perceived from each ear. Phase differences is also visible, where the fifth sound wave here reach the left ear at the same time as the seventh sound wave reach the right. This illustrations is simplified when it comes to wave propagation. 17

3.1 Illustration of the multiple theoretical approaches to the design and development of this prototype and the intermediate steps. 41

3.2 The sound design of the prototype visualized. The *green* section goes from $-108°$ across $0°$ to $108°$ and represent the audio beacon with spatial filtering. The *blue* section in the rear is where the audio beacon is applied a lowpass filter to reduce front/back confusion. The *red* section represent the heading confirmation that goes from $-20°$ to $20°$ and double the audio beacon. 47

4.1 The route the participants walked during testing. It was intended to be traversed from one to seven and back again. The short timespan of the test made us to start from waypoint three for the last two participants. Satellite photo from the Norwegian Mapping Authority (NMA) web service http://www.norgeskart.no/. 58

6.1 Visualized distribution of resolution of a virtual spatial audio display on a horizontal arc in front of the listener. Notice the uneven distribution of resolution in the horisontal and the vertical plane. This illustration is not to scale. 76

6.2 Spectrum analyses of the *Purr.aiff* sound clip used in the Sound Guide prototype. The graph shows that the sound used consisted of a broad range of frequencies. 79

6.3 Examples of two different tracks of route performance. One with an angular bearing with compass sensor (A) and angular bearing based on GPS fixes (B). Illustration from **?**, p 200. 89
6.4 Visualized sound design where rear placement are represented as static *left* or *right* icons to comply with *Guideline 1* and avoid front/back confusion. 93

List of Tables

2.1 Mean times in seconds of all conditions (SD: 95.5-138.4) from **?** 31

3.1 How the audio beacon was filtered in the prototype through the Panorama 5 plugin from Wave Arts Inc. 45

5.1 Results from tests conducted on two participants. *Learning effects* prevents significance testing between the groups. . . . 60

6.1 A collection of results gathered from several studies as presented in **?**, p 39. As these numbers result from different studies that use slightly different methods, the localization blurs is therefore not fully comparable. 79

"We know more about what is involved in getting a man to the moon than about what is involved in getting a blind person across a street."

Emerson Foulke

Chapter 1

Introduction

There are sold about 50% more smartphones than personal computers globally [**??**]. They are recognized as mobile devices with high computing power, a mobile connection and often large graphical screens that offer immersive interaction with applications and entertainment. The introduction of touch sensitive screens and motion sensors has made this interaction more natural than before. Users can now interact with the graphical elements directly using their fingers and through simple motion gestures of the device itself. Graphical user interfaces have matured by several decades of research and has proved its worth in being the face of information systems as we see them today.

Still, there is a large number of people that can not use such graphical interfaces. As of 2010 there was an estimated 285 million people in the world that were visually impaired, where 39 million of these were blind [**?**]. Lack of vision makes auditory interfaces the primary mean of accessing information systems[1]. Voice synthesizing and the growing field of speech recognition have improved in the use of sound as an interactive medium. This has made auditory displays more used even by sighted users that want their hands and eyes free for other tasks.

Auditory interfaces are still slow compared to graphical interfaces. A reason for this is that they are based on one dimension: *time*. Through this single dimension, both content, structure and options are presented in the interface. Let me explain: Imagine reading a web article through voice synthesizing. Information is presented one word after the other in a linear fashion. It might be possible to jump between sections or links, but if you search for some key information in the text you need to listen to it all. The time it takes to read may be altered. Experienced users of auditory interfaces have the rate of talking speed set so high that other people struggle to make the chatter intelligible. However, the medium is still considered linear, in only one dimension. One element before another.

Imagine reading the same article on a graphical screen. Just by a quick glance you will get a lot of structural information. You get a sense of the length of the article, how many sections there are, how many links, if the

[1]Tactile interfaces that rely on Braille letters exist, but since most blind people have not learned to read Braille [**?**] sound can be said to be a more versatile medium.

article has emphasized quotes, tables, figures or pictures and so on. If numbers is what you look for, you will skip directly to the table. If you look for other information, you will probably scan swiftly through the text looking for information of your interest. Without really looking for it, you might recognize the information you look for in the title of a related article presented in the sidebar. When would you have cared to read the sidebar while using an auditory interface?

Information on graphical screens is *spatially* distributed. This gives the possibility of presenting a broad range of information on the same screen side by side at the same time. Sight, as a perception, makes it possible to get structural overview of content, as well as rapid change of attention by the swift motion of our eyes. It can be said that the spatial element of position makes graphical interfaces much more effective than audible interfaces.

This lack of spatial distribution in auditory interfaces is not a limitation of sound itself. Imagine walking into a room blindfolded. Just by listening you will quickly determine the location of people in the room based on the sound of their voices and the reverberation of sound will give you a clue of how large the room is. You will hear an open window at your left side and a female approaching you from you right side just by the sound of her high heals, stopping only a meter away from you. This is examples of our ability to detect *where* sounds are coming from in the environment. *Sound localization* is an attribute of human perception that is quite powerful. Still, the utilization of this ability in design of auditory interfaces of information systems are just in its infancy.

Spatial audio is reproduced sound where the spatial attributes of sound is preserved or controlled. This thesis will look at how this type of audio can be applied in interface design on mobile devices particularity. The novelty of the technology indicate the need to develop general guidelines for the use of spatial audio in interface design. The task of developing a prototype for a smartphone was used to acquire knowledge in the field, both practical and theoretical. It also functions as an example that these types of systems can be built today. It is not the technology that limits us, it is the design and knowledge that have not caught up.

An area where spatial audio can offer many advantages is in navigation. Especially navigation aids for visually impaired pedestrians can benefit from richer auditory interfaces. The limitations of the traditional one-dimensional approach is here apparent and can be seen as highly abstract compared to the use of natural sounds in the environment. The prototype developed in this study try to present route directions in a new way through the use of spatial audio and sensor data. We believe the use of spatial audio has the potential to give richer, more effective, and primarily more intuitive audible interfaces for visually impaired and sighted users than what is available today.

1.1 Research Area

This work is done in the field of *Human Computer Interaction* (HCI) from the larger area of *Informatics*. The applied use of this field is normally referred to as *interaction design*. A *user-centered* approach has been chosen in the interpretation of the results. A range of research areas has been used to accommodate this.

Spatial audio relies on human sound localization through the auditory sense. There will be given an overview of how the human perception system localizes sound sources to understand the limitations of this ability on sound design. Findings that are relevant for our use of spatial audio will also be presented. *Psychoacoustics* is the study of sound perception, where both *cognitive psychology* and human *physiology* are used to explain the perception of sound.

This knowledge has affected how spatial audio is reproduced. An introduction to spatial audio technology is given and an overview of the support of this technology on mobile platforms is presented. The focus on technology is only superficial in this study where their limitations on interaction design is given most interest.

Pedestrian navigation for visually impaired people has been used as a case in this study. *Orientation and mobility* theory will be presented to understand the task and user group relevant for our case. This field concerns how people navigate and move in the environment and is highly relevant for visually impaired pedestrians. An overview of current tools and methods in orientation and mobility will be given.

1.2 Motivation

The personal motivation for this project was given in a lecture on perception in a cognitive psychology class attended four years ago. Sound localization and sound source discrimination was mentioned briefly as abilities of the human perception system. The potential of using sound localization in auditory interfaces became apparent, and it started an interest in the field.

The research field of spatial audio has been around for several decades. It is not large, and few practical use of this technology has reached widespread use. In this work, we want to push the use of spatial audio towards the commercial market, by presenting one domain of applications where the technology has great potential and show that these types of applications can be developed on open smartphone platforms today. We believe spatial audio is a vital tool in the design of new auditory interfaces that are truly intuitive and effective and makes new possibilities for many people. As **?** proclaims: *"Design includes the generation of new possibilities."*[p 170].

Visually impaired people are perhaps the most experienced users of auditory interfaces today. Many people in this user group are used to interact with information systems through audio and should be able to recognize new possibilities and advances in this field and be motivated to use them. However, test have shown that they do not necessary perform better than sighted in using spatial audio displays [**?**].

The disabilities of visually impaired people make them have fundamental needs that technology has the *potential* to meet, but where little is done compared to mainstream technology development. Two of these needs are *accessibility to the environment* and the *independence of movement* [**?**].

Electronic orientation aids for visually impaired pedestrians today normally present directional information as right or left turns. This can be seen as an oversimplification of reality. Experienced users of cane and dogs will encounter situations where such instructions become ambiguous and more precise directional information is needed. The traversal of open spaces, like parking lots or football fields, is an example of one area where traditional directional instructions intended for bounded streets falls short. We believe verbal instructions is not the best choice for directional instructions with high fidelity. Here, spatial audio can provide a more intuitive and more effective method of presentation. It may also reduce the need for time consuming training of users, as **?** arguments: *"since we rely on perceptual rather than symbolic processing, our system requires minimal instructions or training."*[p 334].

Accessibility to the environment involves the ability for people to travel to and from places when they wish to, and the ability to interpret, recognize and understand the environment and their location relative to it. This need is only partially met if we are to compare the abilities sighted people have with respect to those with visually disabilities. Blind pedestrians do not have access to contextual information about the environment like street names, stores, transportation or warning signs without the aid of tools, memory or human assistance. We believe spatial audio displays may contribute in fulfilling the need of accessibility to the environment for visually impaired pedestrians.

Independence of movement is another fundamental need with respect to visually impaired people. Even though aids and techniques have made it possible for many visually impaired to be more or less independent in their everyday life, the majority of blind people are dependent on assistance from others when it comes to traversal of the environment [**?**]. This can be everything from help in learning new routes to being dependent on asking people passing by where the entrance to the store is located. Everyone need some kind of assistance from time to time, but when this need is experienced as a burden and limitation it can decrease the quality of life accordingly. Technological aids have the potential to replace human assistance, or more desirable: make assistance unnecessary.

1.3 Research Questions

In our work, we wanted to learn more about the possible use of spatial audio in interaction design, especially for mobile devices. As a method of doing this, we set ourselves a task to develop a prototype that utilized spatial audio as a vital component of its interface. To learn what is possible for mobile developers today we wanted to make it as a standalone high fidelity prototype on a commercial smartphone. The possible areas of use of spatial audio are wast, but we found visually impaired users and pedestrian navigation to be one area where spatial audio show considerable potential. The research task can then be outlined as:

> **Research Task:** Develope an application for a commercial smartphone platform that utilize spatial audio for directional guidance in route navigation for visually impaired pedestrians.

Much work has been put into the developing period of this prototype, but the end product itself is not meant to be seen as the main contribution of our work. The development of the prototype was based on a broad and multidisciplinary set of theory. Developers of commercial applications that use spatial audio will seldom have enough time to take this massive field into account. As a contribution to interaction design with respect to spatial audio, the following research question was formulated:

> **Research Question 1:** What interaction design guidelines for the use of spatial audio displays in a mobile context can be inferred by theory?

We will here try to formalize the constraints and the recommendations that influence interaction design of *spatial audio displays* with respect to sound localization. Many of these findings will also be valid for desktop applications, but a focus on the mobile context will be given here. The guidelines are most relevant with the use of spatial position as a structural or informative attribute in audible interfaces. The findings are based on use of headphones in the reproduction of spatial audio.

The development and testing of the prototype have revealed a range of findings that is relevant for the design of similar applications. A final presentation of possible improvements of the prototype and its concept of route guidance will be made with the following research question:

> **Research Question 2:** How was the prototype that utilized spatial audio on a smartphone for directional guidance received by visually impaired test users and what can be done to improve the concept in later development?

This question will be answered partly based on the user tests conducted, the knowledge gained from the principles resulting from the first research question and comparisons with other relevant prototypes.

A more detailed description of the scope of this study will be addressed in the next chapter.

1.4 Problem Analysis

In research question 1 where guidelines are inferred, the findings have been mainly based on psychoacoustical theory in combination with interaction design theory. Development and user testing of the prototype has on some aspects influenced these findings, but they are largely based on a theoretical approach. The guidelines should therefore be seen as starting points for the use of spatial audio in interface design, and not as proven limitations or standards.

The approach to research question 2 is a little more complex, and is tied to how the prototype was developed. The user was involved late in the process of development where technology and human perception initially was the main focus. Orientation and mobility theory was then approached *after* the user testing was completed. The field tests were therefore exploratory in nature, with a broad scope and with only a few predefined questions and concerns. The test produced a lot of interesting results, but the broad field of knowledge and the exploratory method of testing made it hard to make firm conclusions based on the gathered material. As with the first research question, the findings should be seen as instructive and not as conclusive.

When technology is concerned, this study will give higher priority to *commercial availability* than to *technological feasibility*. Much research on information systems develops prototypes and 'proof of concept' based on chunky and specialized hardware and software solutions that do not reflect the current technology in the market. This is highly valuable in the context of research, and a similar method could have been used to produce much more advanced spatial audio interfaces in this study as well. Still, new software development inside the technological limitations of the current market is also highly valuable. This choice of focus can be seen as a *critical* choice of the current state of research on navigational aids for disabled. There is a great potential for new highly specialized aids for the disabled using commercial available devices made for 'ordinary' people. Development of specialized hardware is slow and expensive. It also participates in separating disabled people from the community by the tools they use. This study will talk about possible navigational aids for the visually impaired with traditional commercial smartphones in mind to aid in the development of new tools that can reach large markets in a short period of time.

A vital concern when developing electronic navigational aids for visually impaired users is the information available in *geographic information systems*(GIS). This will not be elaborated on in the discussion of navigational aids in this thesis. The focus is more on the method of giving directional information to facilitate orientation, rather than on the implementation and use of information to give a complete picture of the environment.

1.5 Contributions

This work will be of greatest interest for interaction designers that seek to use spatial audio in the design of audible interfaces on mobile devices. It concerns the design and development of such interfaces with respect to the limitations of the human perception and the limitations of the current technology.

A range of preliminary guidelines is formulated based on relevant theory. These guidelines talk about the optimal area for spatial positioning, sound design with respect to spatial audio, how several sources interact with each other, methods of interaction with the display, contextual issues and choice of representation. These guidelines should be seen as a good starting points for design of spatial displays.

Our work has shown that it is possible to develop applications for smartphones that utilize spatial audio for directional guidance. The prototype developed was evaluated with visual impaired test users and will also be reviewed based on the above guidelines. These findings can contribute to the design and development of future electronic orientation aids.

The research that have been conducted during this work has also contributed to other research projects. We have been collaborating with the research project *RHYME* [2] that seek to improve health and life quality for persons with severe disabilities. The *SMUDI* [3] project by *Media Lt* has also been a strong collaborator with respect to interaction with auditory displays. The pilot project *Improved GPS for visually impaired* [4] have shown interest in using this work as part of their coming evaluation.

The knowledge gained in development of the Sound Guide prototype has aided in the development of a music improvisation application for disabled children. This development has been conducted in the pilot project *MIA*, a former student project that are now supported by *IT Funk* [5].

The research conducted have also been featured in the publication *Handikapnytt* [6] that is distributed nationally in Norway by the *Norwegian Association of Disabled*.

1.6 Overview

This thesis can be roughly divided into three parts: the presentation of relevant theory, the description of the prototype developed and the user tests conducted, and the discussion section that uses the former parts in answering the research questions stated.

This work is based on a multidisciplinary field of theory as can be seen in figure 1.1. Human computer interaction and design theory in section 2.1 forms the basis for the design of interfaces and methods of user interaction. Psychoacoustical theory with respect to sound localization is presented in section 2.2, and covers the limitations of human sound perception that is highly relevant in spatial audio displays. The corresponding technology that produce spatial audio is presented and how this technology is supported in today's smartphones are given a brief look in section 2.3. A

[2] http://www.rhyme.no

[3] http://medialt.no/stemmestyring-i-multimodal-dialog-smudi/280.aspx

[4] http://medialt.no/forbedret-gps-for-synshemmede/1149.aspx

[5] http://itfunk.org/docs/prosjekter/MIA_Music_Impro_App.htm

[6] http://www.handikapnytt.no/index.asp?id=77026

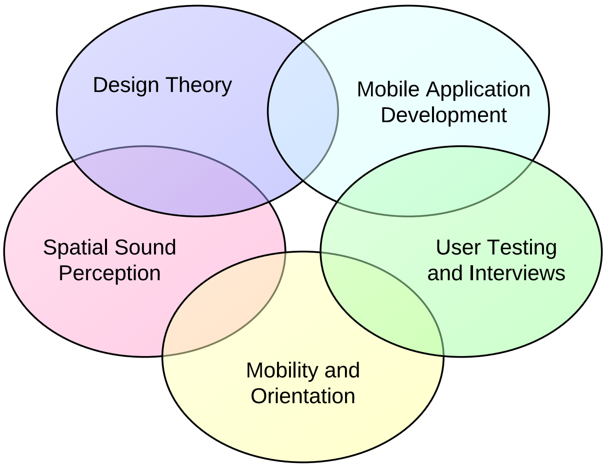

Figure 1.1: A broad range of theory is used in this case in the attempt to give a satisfactory foundation to answer the research questions stated.

selection of applications that use spatial audio is presented in section 2.4. The research prototype called the *UCSB Personal Guidance System* serve as the basis for comparison for the prototype developed in our work. The theory chapter is ended with an exploration of relevant topics in navigation for the visually impaired.

Chapter 3 present the *Sound Guide* prototype developed in this study. Its technical implementation and interface design will be explored. Chapter 4 covers the research methods used in evaluating the prototype and gaining background knowledge in the use context of visual impaired navigation. The results from the use of these methods are presented in chapter 5.

Chapter 6 seeks to answer the research questions stated with respect to the former sections. Interaction design guidelines for the use of spatial audio displays will be formulated based on the presented theory in section 6.1. Section 6.2 will elaborate on the results from the user test and the inferred guidelines to critique to Sound Guide prototype. In chapter 7, we conclude this thesis by attempting to drawing some final conclusions and present topics for future work.

In the appendix the source code of the prototypes developed is made available. The quantitative test results from the test of sound accuracy is presented, and the interview guide and consent form is included.

Chapter 2

Theory

This chapter will cover the relevant theory that is needed in understanding the topics and arguments covered in our attempt to answer the research questions. Spatial audio is presented with respect to how the human perception system localize sound position and how technology can use this ability. The support for this technology on mobile platform is reviewed and relevant applications are presented. In the end of this chapter we will look at visual impaired pedestrians and navigation. The theory presented in this chapter will be used throughout the thesis. We will start by positioning our research in the field of *Human Computer Interaction*.

2.1 Human Computer Interaction

The term *Human Computer Interaction* (HCI) [1] was coined by **?** in the early eighties where the importance of understanding the user was given more priority in the development of computer applications. The field is part of the larger area of *Informatics*, but the center of interest lies at humans as users of information systems and not on the information systems per se.

HCI concerns with the design and development of interactive information systems, and evaluation of it with respect to the user task and the context where it is done. Users at stationary desktop terminals were earlier the main focus of HCI, but the technological development has today expanded the view to include a broader range of technology and larger interplay between users. HCI is relevant in the way we use our MP3-players on the run, to the way we interact with other people when using teleconference systems.

2.1.1 Interaction Design

HCI is an applied science. The role of an *Interaction designer* can be described as *"designing interactive products to support people in their everyday and working lives"* [**?**, p 6] and is now a vital resource in the development

[1] The field was earlier recognized as *man-machine interaction*, but since computers became more popular and both sexes seemed to use them the term *human-computer interaction* was adopted.

of information systems with the user in mind. The process of interaction design can be divided into three main activities: First identifying needs and requirements for the user experience, then develop testable versions of designs that meet these requirements, and last evaluate the designs based on the user experience [**?**]. This process is normally repeated in cycles where the evaluation can reveal new needs and produce new requirements.

The following sections will speak of each of these activities with respect to the design process conducted in this study.

Understanding the User

This work has followed a *user-centered* approach to development. This can be described as focus where *"real users and their goals, not just technology, should be the driving force behind development of a product."* [**?**, p 425]. This involves an early focus on users and tasks where tests on real users forms the basis for evaluation and redesign.

The applied psychology used in HCI is normally based on the information-processing approach where human cognition is thought of as information processing [**???**]. How interfaces are design influence how the task is perceived, understood, performed and learned. If, for example, certain information always is present in the environment, we learn to access the external information through our senses instead of storing the same information to our memory [**?**]. To understand fully why users do as they do, human psychology should be taken into account. Understanding users through the help of cognitive science can be seen too exaggerate with respect to design, but as **?** explain it: *"It is only when we have an understanding of the user at this level that we can understand what makes for successful designs."*[p 9].

A user-centered approach is also beneficial when it comes to the understanding of the task the system is intended to support. The traditional situation can be pictured like this: *"designers often become expert with the* device *they are designing. Users are often expert at the* task *they are trying to perform with the device."* [**?**, p 156]. The users should be used throughout the process of design and development to identify the correct requirements of the system and verify that these requirements have been met. There exist a range of methods in doing this. Some methods relevant for this work will be presented in section 4.2.

The prototype developed during this study involved only one cycle of the design process, where the users were involved late in the development process as evaluators of the prototype and expert users in the task designed for. Preliminary interviews with visually impaired users were planned early in development, but this had to be postponed to be done during the prototype evaluation. The implementation of the prototype will be described *before* the methodological chapter in this thesis since it can not be seen as a direct result of the methods used. We will still considers ourselves as following a user-centered approach in this work, based on how we approach and interpret user feedback of the prototype.

This work put a strong focus on the limits of the human perception with respect to *spatial audio*. Section 2.2 will look closer at the *sound localization* in perception and the technology that has been developed to try to reproduce these effects. The user group approached for this study was visually impaired pedestrians. Orientation and mobility theory will be presented in section 2.5 to uncover the needs of the user group with respect to the task of navigation.

Use of Prototyping

A *prototype* is an envisioned product or feature that is made to answer questions in the design process [**?**]. Prototypes can be made in many forms: From simple paper sketches or cardboard boxes with drawings on them, to functional hardware that emulates the functionality of the finished product closely. These are examples of *low-fidelity* and *high-fidelity* prototypes [**?**]. Prototypes can be tested on users. Users do not need to imagine a product, they can try something that emulates the envisioned design and provide feedback on the immediate experience they have with it. Prototypes are an important tool throughout the design process to make design choices visible and testable.

The development done during this work has followed a somewhat untraditional approach where it early started on development of a high-fidelity prototype . The prototype was set to implement one feature of a navigation aid completely, and concerned less about providing a broad representation of all features such applications should ideally provide. This approach can be called *vertical prototyping* where only certain aspects or features of the finished product are represented in high detail [**?**]. Since the frameworks and programming environment chosen for the prototype ideally could be used in the envisioned product, it share many features with *evolutionary prototyping*, where the prototype itself is developed into the end product [**?**].

Evaluation

When a prototype is made it is presented to users and other people in the design process. Users are normally approached in a formalized matter, where the designer has planned certain tasks they want to test or certain questions they want to get answers on. There exist a broad range of methods for these tasks. Observation and interviews are the most common, but experimental methods can also be used. The choice of method should be selected based on what the designer want to evaluate.

Field studies and interviews are conducted in our work to evaluate the prototypes and concepts made. The section 4.2 will present these methods in more detail and explain why they were chosen.

Theory was given a high priority in the interpretation of results in this project. Psychoacoustics were used to try to describe user experience and behaviour. Orientation and mobility theory was used to give understanding of the context the prototype was tested in. Results from relevant projects were also used to strengthen some findings and critique others. A last field that was applied during the evaluation was *design theory*. No *expert-evaluation* with unrelated experts was used, but we have tried to use *frameworks* and *design principles* as a way of describing the findings related to design aspects.

2.1.2 Frameworks in Design

There are many frameworks, models and theories that are made to inform interaction design and guide research in human computer interaction. Frameworks are *prescriptive* where they give directions in how to understand and approach design tasks and provide concepts for describing them [**?**]. In this section, we will look at two frameworks and seven design principles that are relevant for our case and that will be used in the discussion. These frameworks comes from two highly influencal authors in the field of human computer interaction: Donald A. Norman and Terry Winograd.

Mental Models

An important concept in design is *mental models*. We all form mental models of the world as a way of understanding the things we encounter and interact with. By interpreting the visible structure of things and their perceived actions, we form mental models of how they function and should be operated. **?** call the visible parts of a device the *system image*. The user interpret this system image and construct a mental model of how it functions, called the *user's model*.

A device is normally based on a *design model*, which is a conceptual model on how the designer has envisioned the product to be understood. The design model and the user's model should ideally be the same, but if the system image is not clear and consistent enough to communicate the design model, the user may end up with the wrong conceptual model of its use and functioning. Incorrect mental models lead to confusion and errors in interaction. It is therefore essential for a designer to create a suitable design model that can be communicated through the system image in a way that the user can interpret correctly in the making of their own mental models.

Breakdowns

? describe a way to look at objects we interact with as *invisible* when they function in the way we expect. When an experienced operator is driving a car, the person is not explicitly thinking about how to turn the steering wheel to change the cars trajectory, the driver only thinks about where to go and how to get there. For the drivers sake the steering wheel has become an invisible tool in the task of manoeuvring. The wheel as an object only becomes visible in the case of a *breakdown*. A breakdown in respect of design is not necessary when the car stops working, but rather as *"a situation of non-obviousness, in which the recognition that something is* *missing leads to unconcealing (..) some aspect of the network of tools that we are engaging in using."* [**?**, p 165]. In the case of the driver, the steering wheel would be become visible if the car should continue straight ahead even when the driver turn the wheel to the right. This breakdown would expose the steering wheel as an object because it did not function as expected.

Designers of information systems, or other tools, should anticipate breakdowns and provide possible actions that can be taken when they occur. In the domain of driving, this include not only how the car should be operated and which controls that should be made available, but also provide training and develop skills that help the driver recognize the situation and know how to cope with it.

In the case of the steering wheel that did not function as expected, the cause may be attributed to a slippery road and high speed. Training would help the driver not to question the wheel in such situations, but rather slow down and regain control of the car on the slippery terrain. Some cars are developed to detect when their wheels loose grip on the terrain and automatically administer small adjustments on the speed of each wheel to avoid such situations. It is not possible to avoid all breakdowns by means of design, but in this case both training and automation would help the driver to recognize and cope with the breakdown and continue their activity with less problems or setback.

Norman's Principles of Design

In combination with the two frameworks presented above, we have chosen to reference seven design principles when we talking about design aspects in this thesis. The principles showed here in bold are quoted from **?**, p 188 where they are stated to help transform difficult tasks into simple ones.

Principle 1. Use both knowledge in the world and knowledge in the head. When knowledge to perform a task is available externally, the user does not need to learn how to perform the task by heart to use it. Lets use writing with a keyboard as an example. An experienced typist does not need to look down at the keyboard to use it, but a novice typist has the option of finding the position of each letter based on the label on each key. The knowledge reside in the world, but may be internalized to increase the speed of use.

Principle 2. Simplify the structure of tasks. We do not longer need to draw the shape of each letter while writing, a single click and the computer draws for us. The structure of the task of writing is still the same, only simplified.

Principle 3. Make things visible: bridge the gulfs of Execution and Evaluations. When **?** talk about the *gulfs of execution and evaluations* he thinks about the importance of letting the user see the results of their executions as a mean to evaluate them. When a typist hit a key on the keyboard, the corresponding letter will usually appear on the screen. If the letter is not what the typist intended, he/she may notice it and correct it before continuing.

Principle 4. Get the mappings right. The backspace key is normally labelled with an arrow pointing to the left. This can be said to provide a good mapping between what the user intent to do and the effect the action have on the system: The cursor jumps back in the direction of the arrow and removes one letter. If the same keyboard is used to write in a language with a script that goes from right to left, the same mapping will no longer be right, and confusion may occur.

Principle 5. Exploit the power of constraints, both natural and artificial. Constraints in design can be used to only allow the correct actions to be taken. Without removing the experience of control from the user, the interface should present only those actions that a user should need to perform at that step in the task.

Principle 6. Design for error. The designer should assume that any error that *can* be made, *will* sooner or later be made. By planning to prevent errors, or making it easy to recover from them, the interface becomes more usable. The backspace key mentioned earlier is a good example of how mistakes in writing is easily recoverable.

Principle 7. When all else fails, standardize. Some times there is no obvious mapping or design choice that makes an action, outcome or layout self explanatory. In such situations standardization should be approached. This enable users to learn the standard once and then use it effectively each time they encounter it. Since keyboards now mostly follow national and international standards it is possible to use a new keyboard as effectively as an old one, without learning a new arbitrary method of operation from scratch.

These seven principles will be referenced in relation to design choices later in this thesis.

2.1.3 Auditory Displays

Even though graphical user interfaces dominate in human computer interaction, there is also a variety of *auditory displays* that use sound to convey content, structure and options in the user interface. Synthesized voices, recorded speech or abstract sounds are normally used in such interfaces. *Audio icons* are abstract sounds that are used to represent certain elements in an auditory interface - much like visual icons are used in graphical user interfaces. This can be musical tones, recordings, or synthesized sounds that are short, distinct and recognizable.

Auditory displays can be interacted with through traditional keyboards, or other key based input mechanisms. Mouse interaction is seldom used. Voice recognition has the last years matured and start to become more and more natural as a way of interacting with auditory interfaces [**?**].

Examples of Auditory Displays

Many automatic reception systems reachable by phone use auditory displays to guide the caller to the correct person, desk or information. Most such systems present the users options through reading out each option one after the other. To select the option they want, users are instructed to type the corresponding number key on the phone. Some systems use voice recognition to let the caller repeat the option they want, or simple explain what they want in their own words.

GPS navigation aids used in cars is usually multimodal, in the sense that they use an auditory display together with a graphical display. These auditory displays used synthesized voices or speech recordings to present each deviation from the road the driver has to make to reach the specified destination. This information is usually given in short utterances like *"Turn right after 100 meters"*, and are presented in advanced of each crossing. The same information can be found on the screen for verification and clarification, but auditory information is provided to let the driver be able to keep their full attention at the road and the task of driving if needed.

Screen readers are applications that translate traditional graphical user interfaces into auditory displays. They are intended as alternatives for blind and visually impaired users that are not able to use traditional graphical interfaces. Screen readers are found on computers and mobile devices and function by using synthesized voice to list information and option that is present on the screen and provide methods of interaction to accommodate this.

Spatial Attributes in Auditory Displays

As argued for in the introduction, auditory displays are usually one dimensional where content, structure and options are presented serially, one after the other. This can be likened to reading a restaurant menu through a straw where the attention can only be given one word at a time. These constraints lead to rigid and slow interfaces that struggle to compete with the graphical user interfaces in effectiveness and usability. Auditory displays that are able to display richer interfaces of content, structure and options are needed.

> *"Augmenting a sound system with a spatial attribute opens new dimensions for audio; spatial sound is a rich audio analog of three-dimensional graphics."* [**?**, p 320]

Spatial audio has in the last decades captured the attention of researchers as means of designing better auditory displays. If sound can be given a spatial attribute, a position in relation to other sounds, it can be used to convey structure, information or options beyond the semantic meaning of the sound itself.

Spatial audio has been used experimentally in a variety of ways to explore these possibilities. **?** designed and experimental auditory display that presented menu structures with elements positioned in a semi-circle around the user. Improved screen reader was attempted designed by **?** with using spatial positioning to mimic the visual distribution of elements in a graphical interface. **?** developed a wearable system for voice- and text-based messaging that used spatial position of audio to indicate which time of day the message was received. A concept for an auditory web browser that used spatial positioning of audio icons to convey document structure was designed by **?**.

These are only some examples from a wide variety of studies that use spatial audio to design new forms of auditory displays. To be able to design such systems, we need to know what spatial audio is and how localization is achieved by human perception. The next section will look into these topics to provide theoretical grounding for later chapters.

2.2 Spatial Audio

> *"Localization" is the law or rule by which the location of an auditory event (e.g., its direction or distance) is related to a specific attribute or attributes of a sound event, or of another event that is in some way correlated with the auditory event.* [**?**, p 37]

If somebody try to get your attention by calling out loud, their voice will carry a lot of different types of information. Their choice of words may convey who the intended receiver is and the purpose of the interruption. If it is a familiar voice you will probably recognize who it is just by listening. If the voice is not familiar you may recognize where the person originates from based on their accent and their gender based on their tone of voice. Their tone of voice may also give away their emotional state and the rate of urgency in their request for attention. In addition to the knowledge acquired above you will probably, without thinking about it, turn your head in the direction the person is calling from. Without looking you will also have an impression of how how far away from you the person who called are.

The location of the sound source with respect to the listener was in the case above identified through the sound itself. Human hearing is able to detect the position of an sound source with a decent accuracy based on how it is perceived. Sound representations that convey this type of spatial positioning will be referred to as *spatial audio* in this paper. Some papers referred to it as *binaural audio* [see **??**], *virtual surround* [see **?**] or *3D audio* [see **???????**], but spatial audio seems to be the term adopted by most of the scientific community on the subject and will therefore be used.

Spatial sound seems to concern two broad categories of sound events that may be characterized as either *source* or *environment* [**?**]. Sources may be described as distinct and localized audio that is perceived to emit from a specific location in space, as in the example above. Spatial audio in the

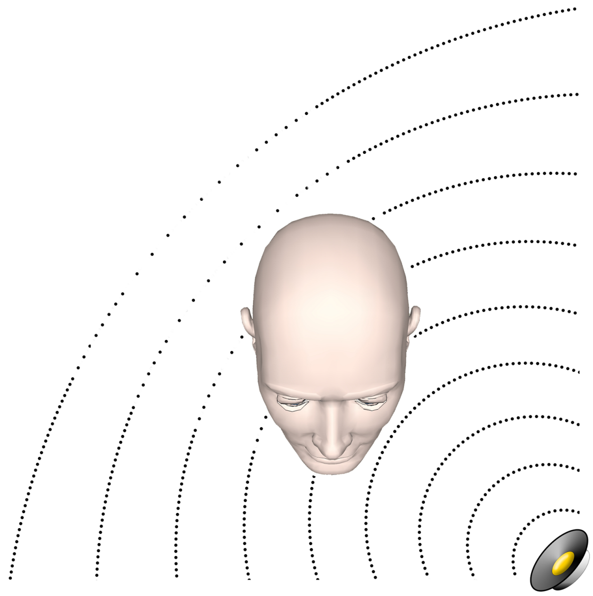

Figure 2.1: This illustration show how occlusions by the head can lead to intensity differences between sounds perceived from each ear. Phase differences is also visible, where the fifth sound wave here reach the left ear at the same time as the seventh sound wave reach the right. This illustrations is simplified when it comes to wave propagation.

environment category concern more ambient audio like room reflections and background noise that give a more general impression of the space around the listener. This paper will focus on the source characteristics of spatial audio, and use the term with that aspect in mind.

2.2.1 Psychoacoustical Cues in Sound Localization

Sound localization is a complex process where the brain use a range of cues to identify the spatial position of sounds sources [**?**]. In sound localization it is common to draw a distinction between localization in the horizontal and the vertical dimension. Localization in the horizontal plane, defined by the tip of the nose and the two ear canals, is largely based on *interaural* differences between the two ears [**?**]. This is sometimes referenced as *binaural cues*, or *binaural hearing*.

Horizontal Localization

Humans have two ears located in a fixed position on each side of the head. If a sound is presented from the side it will be shadowed by the listeners head in reaching the far ear. That will make the sound appear stronger in intensity on the side it originates from (see figure 2.1). The interaural differences of the two ears in sound pressure level will indicate which side the sound originates from and contribute in localization of the source.

This method of localization holds true for sounds of high pitch of tone. For frequencies below 1000 Hz the sound wavelength will be larger than the head and the sound pressure level will no longer give sufficient cues to determine localization.

Lord Rayleigh, a British physicist, discovered that the difference in phases at the two ears had to account for localization of sound with low pitch [**?**]. If a sound is presented directly from the side it will reach the closest ear first and arrive at the far ear approximately one millisecond later [**?**]. The amount of phase differences will decrease as the angle from the front decrease and be zero when the sound is directly ahead of the listener. The brain detect this differences and use this to determine sound source position in the horizontal plane. The sensitivity of phase differences declined with increased frequencies and is not present above 1600 Hz [**?**].

Interaural differences in phases determine localization in the horizontal plane for low frequencies, while intensity differences determine localization for high frequencies. This has been known as the *duplex theory* of sound localization [**?**].

Vertical Localization

With respect to sound localization in the vertical dimension interaural differences falls short. Vertical movement in the medial plane will not produce any interaural differences when the ears and heads are symmetrical. Still: humans are able to detect if a sound is coming from above or below. The main cues used in vertical localization of sound seems to be something called ”spectral shape cues” [**?**]. The shape of *pinna*, the external part of the ear, produce subtle changes in the spectrum of the sound based on the direction and distance it originates from [**?**]. When sound enter through the ear canal it is slightly altered by reflections, shadowing and resonance caused by the external ear. This cues appear as peaks and notches in the spectrum at certain frequencies that the sensory register manage to detect and interpret as localization cues[2]. This type of vertical localization seems to rely on broadband sound stimuli where frequencies above 4000 Hz is most important [**?**]. The effects of spectral shape cues on localization in the horizontal plane are much discussed, but it does seem to aid in reducing front/back confusion. Reflections from the shoulders and torso may also aid in sound localization in a similar matter, but the use of these cues seem to differ considerably between individuals [**?**].

[2] Vertical localization seems to be almost as accurate with one ear as with two. In auditory interface design it could be possible to utilize this aspect of spatial audio in applications that use headphones with a single earpiece. Handsfree Bluetooth sets that present notifications from above while the conversation is presented from the side, may be a possible application.

Distance Localization

How we localize the distance of sound sources has been less researched and is not as accurate as directional localization. Distance can be judged with some accuracy based on the sound intensity if the source intensity already is known [???]. In that sense it will be more easy to place the distance on speech than on an abstract tone where the level of intensity of the source is unknown. If the sound is emitted in free-field, an environment without reflections, the sound volume will drop by 6 dB for every doubling in distance from the source and air absorption will give less high frequency response [??].

In reflective environments, where walls and other objects may reflect soundwaves, the direct sound perceived is followed by reflections and reverberation. While reflections are more clear echoes, reverberation consist of numerous reflections of reflections that may be likened to a diffuse 'shadow' of the direct sound signal. Reflections and reverberation provide cues that greatly help in distance perception, but may in some cases make it more difficult to localize the direction of the sound source [???].

Front/Back Confusion

A common error in localization sound sources is front/back confusion. Is the sound source straight ahead or behind the listener? This type of errors is linked with the frequency of the sound. Localization of pure tones of around 1500 to 3000 Hz gives the largest amount of localization errors [?]. This may be explained by using the duplex theory where the frequency range is to high for detection of phase differences and to long in wavelength to provide enough intensity differences [?]. The slight forward facing of the ears will also reduce the intensity of higher frequencies when the sound is coming from the rear [?]. The use of broadband sound stimuli, sound that consist of a broad spectre of frequencies, reduce the amount of front/back confusion and is therefore preferable over pure tones in facilitating sound localization. Another important aid in removing front/back confusion is active movement, as described in the following section.

Head Movement

The psychoacoustical cues in sound localization that are discussed above is all passive in the way that they deal with interpretation of the sounds that reach the two ears. A much discussed active clue in sound localization is head movement. When a subject is asked to determine the location of a sound source it is common of the subjects to spontaneously turn their head toward the sound [??]. This is the primary way of reducing front/back confusion and is known to supersede the spectral shape cues in the task [??]. It is not known if head movements aid in localization by facilitating comparison of sound perceptions from two head positions, or if it is the facing of the source itself that improve upon localization [?]. The increase in localization accuracy given by head movement is still minor compared to the other cues.

Cues Combined

It is much discussion concerning which cues that have the greatest influence on sound localization. To summarize the view of **?** localization is determined by a combination of cues where the different cues provide information on different aspects of localization. Interaural differences provide the strongest cues in horizontal localization of sound sources, while spectral shape cues determine the localization in the vertical dimension. Front/back ambiguities is primary resolved by active head movements or secondary by spectral shape cues. Since each clue may provide ambiguous information the full power of sound localization is only present with a combined use of several cues [**?**].

2.2.2 The Cocktail Party Effect

A perceptual by-product of using spatial audio is the aspect of sound source discrimination. Binaural hearing gives the listener the ability to pay attention to one audio stream even though there are several audio streams that are played at the same time. The spatial position of each audio stream helps the brain distinguish between them and focus its attention on one of them [**?**]. This phenomena is in psychology often referred to as the *Cocktail Party Effect* [**?????**]. Named after our ability to pay attention to one person talking in a room full of other concurring conversations. Spatial position is one of many factors that aid in such sound source discrimination. Other factors are intensity differences, pitch differences, differential filtering and degree of synchrony [**?**].

This effect of sound localization makes it possible to design auditory displays that present verbal information from several sources at the same time. **?** used spatial audio to present seven competing talkers in different configurations around a listener during an experiment. **?** designed a teleconferencing system that enabled the listener to position other talkers in a virtual environment by hand gestures. The aspect of sound source discrimination is quite influential in the design of spatial audio displays, but will be given less focus in this thesis.

2.2.3 Stereo Audio

Much audio recordings produced today consist of two channels played together as one stereo track. In music production these two channels are used to distribute the sound field to produce more interesting and aesthetically pleasing listening experiences. When played back over speakers or headphones this give the impression of crude spatial placement of different sound sources. Guitars may be positioned to the right, drums to the left while the vocal originate from the middle. These types of spatial placement is only achieved through intensity differences between the left and right channel. No other binaural cues is normally present and the impression of spatial placement is therefore quite crude compared to natural sound localization in the environment. Listeners often report the perceived sound source position of traditional stereo tracks presented over headphones to originate from inside their head right between the ears [**??**]. Panning between stereo channels may be used in auditory displays, but balanced representations seems to be more common.

2.2.4 Binaural Recordings

Since the psychoacoustical cues used in sound localization are given by eternal characteristics as head position and shape of the ear canal and torso, it is possible to record and reproduce it. By using small microphones placed inside the ears of a human head, or a dummy head, the stereo recording will contain the sound intensity, phase differences and spectral changes used in sound localization. When played back over headphones, this types of recordings may reproduce the auditory space of the recording with great realism [**?**]. The effect depends on the shape of the head and the ears on the recorded head compared to the listener's, where it ideally should be the same head. The binaural recording would not give the desired effect when played over loudspeakers and is therefore reserved for headphone listening. There was some early attempt in the 1970s to introduce binaural recordings of music in the broadcast and recording industry [**?**]. The attempts failed however due to the problem with individual differences (many listeners experienced "forward-to-backward confusion" when listening to binaural recordings) and the limitations of headphones as the only mode of playback.

2.2.5 Head Related Transfer Functions

> *It becomes apparent at this point that auditory scenarios can be generated even without any recordings of real-head or dummy-head signals in real scenarios.* [**?**, p 375]

If we know all the cues the auditory system use in sound localization we should also be able to emulate them. Electronic filters that try to replicate these cues are called *Head Related Transfer Functions* (HRTF). When listening to a simple unfiltered recording through headphones it will appear that the sound originates from a place inside ones head, right between the ears [**?**]. If the same recording is preprocessed in a HRTF that manage to emulate the psychoacoustical cues discussed above, the sound may be virtually placed outside of your head at a chosen position. The sound will appear to originate from that position as if it was placed a speaker there playing the original recording. This is often called an 'out of the head' experience. All positions around a head may be simulated, but with different accuracy. There exists HRTFs that are intended for loudspeaker playback, but this paper will primarily focus on headphones listening with respect to spatial audio.

There is a range of different HRTFs made. They are normally based on large samples of binaural recordings, with dummy heads or real subjects, where the source audio and the position is known. Some HRTFs only emulate intensity- and phase-differences, while others incorporate emulations of spectral shape cues as well. As with fingerprints, there is large differences in the size and shape of peoples head, ears and torso [**??**]. While one HRTF may give a rich 'out of the head' experience with accurate sound localization for one individual, it may give a poor reproduction for another. Ideally a HRTF should be tailored to each individual, based on their head and ear shape. Much research is done on producing general HRTFs that works well across a large number of people. A mere averaging of individual HRTFs is not sufficient, but an identification of the most important cues and how to utilize them is needed [**?**]. While some studies indicate the possibility of generalization, it is still best with individual HRTFs [**??**].

There are still studies that indicate large plasticity in the brain regions that is used in sound localization, where the brain can learn to adapt to large differences in the auditory cues and reach a localization accuracy close to the original set [**??**]. Given a rich and responsive environment where the localization accuracy may be trained, it should be possible to grow accustomed to general HRTFs and improve localization.

2.3 Mobile Hardware and Software Support

The term for spatial audio adopted in the consumer market is *3D audio*. What this represent is often vague and imprecise and should be seen as an umbrella term used in marketing. Some specialized features based on limited spatial audio techniques is becoming more common. *Stereo widening* is a feature for headphones playback where simple digital audio filters are applied to the audio output to give the sense of a broader soundscape. This is sometimes refered to as a 3D audio feature. *Virtual surround* is a feature that use a simplified HRTF filter set to emulate surround sound over headphones. This technique position 8 virtual speakers in fixed positions around the listener locked to the head position. This is also a common feature under the umbrella term 3D audio. When the term spatial audio is used in this section it represents the use of complete HRTF profiles where no technological limitations is present in the positioning of sound sources.

When the use of spatial audio in interface design in applications for mobile devices is approached, the device support for the technology has to be taken into account. Seen from a developers standpoint, there are very shallow support of this technology on today's smartphone platforms. Even though major manufacturers have done a lot of research and development, very little practical support for spatial audio has reached commercial platforms to enable real use of these technologies [**?**].

To get an overview of the current support for spatial audio technology on commercial devices today, both hardware and software frameworks had to be reviewed. Even though it is possible to develop spatial audio support on the software level (see 2.4), the high processing demands of the transfer functions involved in HRTF filtering seems to speak in favour of dedicated hardware support for these features. A short review of the major chipset manufactures for mobile devices may indicate the current possibilities, while a short review of how software frameworks are prepared for these features may indicate how it may be used by third party developers.

2.3.1 Hardware Support

The company *ST-Ericsson* is a major supplier of chipsets for mobile operators and device manufacturers. A chipset is normally a centralized hardware that provide processing power, graphics, audio and access to all other hardware components like wireless and camera. Their processor *AP9500* in the *NovaThor U9500* platform has good support for spatial audio filtering. The processor is stated to support stereo widening, 3D positioning, Doppler effects, reverberation, virtual surround and several other advanced audio features[3]. only one smartphone (*HTC Z710t*) that is built with this specific framework[4] were found, but specifications are limited so there may be more frameworks from ST-Ericsson that supports these features.

Qualcomm is another major supplier of chipsets, and present their *Snapdragon S4* processor as their leading product for mobile devices. This chipsets delivers all multimedia and processing capabilities to the very successful smartphone *Samsung Galaxy SIII* and a range of other devices on the market[5]. When the specification is given a closer look the only stated 3D audio support is virtual surround sound following the *Dolby 7.1 Surround Sound* specifications[6].

Texas Instruments is another major chipset supplier that has been in the business a long time. Little is said about the audio support in their *OMAP 5 Platform* [7] that is used on a wide range of smartphones, but they sell another chip that is quite interesting. The *LM48903 Stereo spatial audio array* enables positioning of several sound sources through HRTF filtering. The product does seem to focus on speaker systems, and advanced features like distance attenuation and room reverberation is not mentioned. The existence of commercial smartphones utilizing this chip is not mentioned.

There are several manufactures of similar specialized chips that expand the audio capabilities of existing chipsets on mobile devices. *QSound* is the developer of a digital audio engine called *microQ* that can position

[3] `http://www.stericsson.com/developers/DM00030004_AP9500_reference_manual_rev1.pdf` (Accessed 20. June 2012)

[4] `http://www.stericsson.com/press_releases/HTC_sensation.jsp` (Accessed 20. June 2012)

[5] `http://www.qualcomm.com/snapdragon/devices/finder?processor=4` (Accessed 20. June 2012)

[6] `http://www.qualcomm.com/media/documents/snapdragon-s4-processors-overview` (Accessed 20. June 2012)

[7] `http://www.ti.com/general/docs/wtbu/wtbuproductcontent.tsp?templateId=6123&navigationId=12863&contentId=103102` (Accessed 20. June 2012)

multiple sound sources with spatial filtering and may also produce room simulations with echo and reverberation[8]. It is not stated if there exist commercial devices on the market that use the product.

2.3.2 Software Support

While the ability to use real time spatial audio filtering is mostly dependent on hardware support, the software framework that applications reside in has also much to say. It does not help if a device has hardware support for HRTF filtering if the software framework prevent third-party applications to access these features. All third-party applications made for smartphones have to access a device hardware features through a framework or *API* (Application Programming Interface). For this review the most used frameworks have been reviewed for spatial audio support.

OpenSL ES on Google Android

With the release of the mobile operating system *Google Android 2.3* in December 2010, it was know that the platforms native audio APIs should be based on the *OpenSL ES 1.01* Standard. This audio API for embedded systems is developed by the *Khronos Group* and have 3D positional audio as one of its main features [9]. Using this API one should be able to control the position of audio sources with just a few lines of code. This was interpret by some to mean that the Android platform now had native support for spatial audio. Careful reading of the native audio documentation for Android NDK revealed that the OpenSL ES 1.01 standard is only partially supported on Android and 3D audio is not mentioned on the feature list [10]. Stereo panning and audio effects like room reverberation is implemented, but spatial positioning of audio sources is not.

OpenAl on Apple iOS

OpenAL is another, more widely known cross-platform audio API that states to support 3d audio. This API is implemented in the platform *iOS* which is found in *Apple iPhone* and is also portable to Android. The API handles listener and source objects and the placement and even velocity of these in a virtual environment. Testing reveal that even though the library states its primary purpose is to allow placement of audio sources in a 3D space around the listener [**?**], it leaves the actual rendering of spatial audio position to the implementation environment . This means that if HRTF's is not supported on the hardware level, positioning through the OpenAL library will only produce stereo-panning, volume attenuation and doppler effects. In game development these features will still be beneficial

[8] `http://www.qsound.com/technology/3d-audio.htm` (Accessed 20. June 2012)

[9] http://www.khronos.org/news/press/khronos-opensl-es-1-1-specification-for-stereo-3d-audio-mobile-os (Accessed 19. June 2012)

[10] `http://mobilepearls.com/labs/native-android-api/opensles` (Accessed 19. June 2012)

for producing audio effects that accommodate and improve the experience of the graphical virtual environment on the screen.

XNA Game Studio on Microsoft Windows Phone

The relatively new mobile platform Windows Phone follows in the same path as the other players. Windows Phones use the XNA Game Studio 4.0 framework for audio control and does also state to support 3D audio [**?**]. While the position of the listener and source may be positioned, the spatial effect is only implemented as a mix between the stereo output [11].

Web Audio API for HTML5

HTML5 is by many stated to be the framework of choice for cross platform applications in the coming years. Its audio support has been very limited, but a new proposed standard called *Web Audio API*[12] has the potential to change that. The API is only a draft and *Google Chrome* and *Apple Safari* is the only web browsers for desktop computers that currently support it[13]. The API offer positioning and spatializing of multiple sound sources in a virtual environment, and support several other advanced features like *occlusion*, *doppler shifts* and *room simulations*[14]. It is not specified if the HRTF filters are a part of the framework itself or if it utilize the underlying hardware support of the system for these calculations.

2.3.3 Commercial Devices with 3D Audio Support

A specification for the *Java ME* platform called *Advanced Multimedia Supplement* (AMMS)[15] have 3D audio as one of its main capabilities. Most devices running the Java ME platform or the *Symbian* operating system does still not support this specification fully. The *Nokia N95* and the *Nokia N97* are two of only a few commercial devices that actually support the full specifications with limited 3D audio capabilities [**??**].

The smartphone *LG Optimus 2X*[16] was launched with virtual surround sound capabilities, as was the smaller phone *LG KM335*[17]. This did not enable full positioning of audio in a virtual environment, but rather

[11] `http://msdn.microsoft.com/en-us/library/bb447687.aspx` (Accessed 19. June 2012)

[12] `https://dvcs.w3.org/hg/audio/raw-file/tip/webaudio/specification.html` (Accessed 19. June 2012)

[13] `http://chromium.googlecode.com/svn/trunk/samples/audio/index.html` (Accessed 19. June 2012)

[14] *Occlusion* is the effects on the sound obstacles generate. *Doppler shifts* is experienced when sound sources of high velocity move past the listener and a decrease in pitch is audible. *Room simulations* is used to model realistic sound reflections that accommodate the shape and size of the room. These effects will not be discussed further in this work.

[15] `http://www.jcp.org/en/jsr/detail?id=234` (Accessed 20. June 2012)

[16] http://www.am3d.com/home-english/news/archive/new-lg-optimus-2x-has-virtual-surround-sound-developed-by-am3d-(1).aspx (Accessed 20. June 2012)

[17] `http://www.lg.com/ae/mobile-phones/all-mobile-phones/LG-camera-phone-KM335.jsp` (Accessed 20. June 2012)

enabled an enhanced playback of surround tracks with 8 virtual positioned loudspeakers in fixed positions. The smartphone *LG Optimus 3D* was said to support 3D audio capabilities from *DTS*[18], but even though we had the ability to develop for this phone no such support was found. As with the other LG phones, virtual surround may be the only 3D audio support they provide.

The previously mentioned smartphone HTC Z710t is built with hardware that supports spatial audio, but none of these features is found mentioned in press releases and technical specifications of the phone. We have not had the chance to test this device, but it could be that even though the hardware support these features, the software framework that access them does not.

2.4 Projects Utilizing Spatial Audio

There exist many research projects that is of relevance for the work conducted in this study. Many of these will be referenced in the discussion chapter and used to explain various findings and concerns in the different research questions. This section will first focus on the few commercial applications and projects that has been made utilizing spatial audio as a main component. And secondly it will present a research project conducted at the University in California Santa Barbara that is highly relevant in relation to our work.

2.4.1 Commercial Applications

There exist only a few commercial applications that utilize spatial audio as an active part of the interface. Some key applications and projects most relevant for this study will be reviewed here[19].

Papa Sangre

The british company *Somtin'else* has develop a game for the Apple iPhone called *Papa Sangre*[20], that utilize spatial audio as one of its main components. Even though no hardware support for spatial audio exist on the Apple iPhone devices, the company managed to develop a real-time 3D audio engine on the software level. This enable them to distribute a

[18] `http://www.dts.com/corporate/press-releases/2011/08/3d-smartphone-featuring-tri-dual.aspx` (Accessed 20. June 2012)

[19] Other commercial applications and projects that can be worth mentioned are: *Aves* is an iPhone game that is said to use real-time 3D audio, but little information is given on the binaural technology used and preliminary tests of the game audio do not manage to impress us (`http://itunes.apple.com/us/app/aves/id321295493?mt=8`). *Dolby Voice* is an audio conference system that utilize spatial audio to segregate speakers in a virtual soundscape (`http://www.dolby.com/us/en/professional/technology/voice/dolby-voice.html`). *InvisiBall* is a project that state to let people play virtual tennis through spatial positioned sound from a speaker setup (`http://hakanlidbo.com/archives/2075`).

[20] `http://www.papasangre.com/` (Accessed 21. June 2012)

game through *Apple App Store* that include its own spatial audio capabilities without more requirements.

The game is played with headphones and present an immersive virtual world through audio as the only medium. A turning wheel on the screen help the player turn around in the virtual world and a sequence of tapping is used to move forward. The aim of the game is to rescue someone in grave danger in a dark dream land. This is done by traversing levels where the player is set to locate and walk to different sounding beacons in the virtual environment. While navigating the virtual audio environment, the player is also hunted by creatures that will attack if they get too close. All objects in the environment is presented through spatial audio where directions and distances of beacons and other creatures are possible to determine by binaural listening.

To accomplish this the company developed their own 3D audio engine that were compiled together with the application to be run in an iOS environment. The engine was written in C++ and communicated directly with audio components in the Cocoa framework without using the OpenAL library [Paul Bennon, Chief Creative Officer in Somethin'Else, personal communication in November 2011]. The developers in Somethin'else had a hard time developing the engine where their main challenges were limited processor power of the device, interrupts and threading. Still, they were able to use fully interchangeable HRTF profiles without simplifications, position up to 5 simultaneous sound sources (on an Apple iPhone 3GS) and support procedural reverberation that take room characteristics into account. Sound source position in the median plane is not supported through spectral cues, but the impression of medial position is accomplished through sound design and dialogue. If the listener expect the sound to originate from above, it will also be perceived in that way: 'Look up, a helicopter is flying over you'.

The company state to be the first to accomplish making a fully functional real-time 3D audio engine for a mobile device[21] when published in late 2010[22]. They plan to make the engine available for third party developers, which may benefit smartphone developers in other fields than the gaming-industry. It will be very interesting to see what type of impact the release of Somethin'else's 3D audio engine will have on the developer community.

Oh Music Where Art Thou

The Android application *OMWAT* [23] was a result from the *Oh Music Where Art Thou* project done by students at the Technical University of Eindhoven in the Netherlands [**?**]. The application is designed to be used while bicycling to give turn-by-turn route guidance. Instead of providing visual

[21] `http://www.papasangre.com/blog/` (Accessed 20. June 2012)

[22] A limited 3D audio effect of stereo widening was achieved on a *Nokia N80* in 2007, developed on the software level with no hardware support [**?**], but this had far from the same possibilities as the engine developed by Somethin'else for the Apple iPhone.

[23] `http://www.usinet.nl/omwat/` (Accessed 21. June 2012)

cues on the screen, the device is made to present direction by spatial positioning of music playback between the left and right channel. The user can choose a destination on the map and a self made playlist with music. The application then calculate the shortest route and while playing music throughout the trip all route descriptions is given through panning of the music. If the user should take a right turn in the next intersection, the music will pan to the right to indicate the desired direction.

Orientation was determined through the device GPS-positioning and compass sensor (magnetometer). Since the orientation of the body is more or less fixed when biking, the smartphone was attached to the headphones which enabled interaction with the audible display through head rotation. A basic form of spatial positioning of music was accomplished using simple stereo channel mixing. This can still produce clear distinctions between left and right and can also indicate more subtle directions when balanced. The volume at the left ear was at maximum when the target was located directly to the left of the user, this was decreased if the user turned towards the target and was at the right maximum when the target was located to the right. No special indication of front and back is given. The loudness of the music was used to indicate distance to the target, where the music grew louder as the user got nearer to the target.

This system proved quite usable in user testing where two users managed to follow a given route on about 1 km and identify the main target [**?**]. We have done preliminary tests of the commercial version (OMWAT). While it proved a bit unstable in giving directions while walking, it gave stable and correct directions while bicycling. This could be based on the advantage of speed when determining heading through GPS-positioning.

The Sound of Football

The technology company *Society46* in Sweden has done a project named *The Sound of Football* [24] that served as a great inspiration for the technical implementation of the prototype in our study. The application developed has never been published and the project can be categorized as a commercial research project.

The motivation for the project was to develop a way for visually impaired people to play football by the aid of an augmented sound landscape. With headphones and an Apple iPhone attached to the top of the head, a virtual football field was given through spatial sound where the field, the ball, the goal and all the other players were identified through different sound icons with spatial position. This enabled them, as a marketing scheme for *Pepsi*, to host a fight between visually impaired players and blindfolded professional football players. Only edited video is available from the project, so little can be determined by the results, but the interviews given and the clips from the game indicate a surprisingly functional setup. Some players reported it to be very difficult to determine

[24] `http://society46.com/the-sound-of-football-pepsi-refresh/`

where the ball was at all times. The video show players at times uncertain, but still able to locate the ball and kick it towards the direction of the goal.

Through the use of *Tracab* [25] tracking system the position of the ball and all the players on the pitch was tracked. This data was then used to determine the players location in relation to each other and the ball. The orientation of each player was given by combining the iPhone's gyroscope and compass sensor (magnetometer). The virtual soundscape was then created by spatial positioning of sound icons that represented other players, the pitch and the ball with respect to the listener. If the player turned around, the virtual soundscape would turn to represent the new orientation. If the player came closer to different elements, the sound would grow louder to indicate the distance to it and the rate of repetitions would grow faster. The players could hear the ball at all times, but the sound of the edges of the pitch was only audible once the player got too close.

The spatial position of sound was achieved by using pre-filtered audio clips. Instead of trying to develop real time spatial audio filtering (as with the Papa Sangre game) each type of sound icon was processed with spatial filtering on a computer to emulate eight different locations in a circle around the listener [26]. These audio clips was then loaded into the iPhone application. By using the proprietary audio library *FMOD*™ these different sets of audio clips was played dependent on the location and distance of the object. If for instance the ball is located 22° to the left of the player, the audio played back will be a mix of the audio clip filtered to emulate 45° and the audio clip filtered to emulate straight ahead. The distance of the object then determined the volume of the playback. This setup makes it possible for the iPhone to present real time position of multiple objects around a listener that is able to move and interact with the environment.

2.4.2 The UCSB Personal Guidance System

A research team at University in California Santa Barbara has in a period of several years developed a personal guidance system that utilize spatial audio in route navigation for visually impaired users [**?????**]. The prototype they developed has many similarities with the Sound Guide prototype developed during our study. We were not aware of this research during development and user testing, but since the prototypes share so much common features much room will be given to compare the two systems later in this thesis.

Loomis was in 1985 one of the first that proposed using GPS to help blind and visually impaired people navigate in unfamiliar territory [**?**]. Instead of using synthesized speech to communicate route guidance, a virtual sound display could provide more directly perceptual information of the environment.

[25] `http://www.tracab.com/`

[26] A larger set of directions was attempted, but performance issues limited the number to eight

The long term goal of the research group is to *"contribute to the development of a portable, self-contained system that will allow visually impaired individuals to travel through familiar and unfamiliar environments without the assistance of guides."* [**?**, p 196].

The prototype developed was not intended to replace mobility aids such as white canes and guide dogs, but rather help in the more general navigation and orientation challenges. The prototype was developed with a user-centered approach where iterative development was guided by test results with visually impaired users. The system that was used in **?** will here be presented.

Concept

The system was designed to guide participants through a route that consisted of several waypoints positioned at corners and intersections. The prototype was used to guide the visually impaired participant in the direction of the next waypoint. When the position of a waypoint was reached, the next waypoint in line was presented. A waypoint was considered reached when the user was within 2.1 meters of its position.

The orientation of the waypoint with respect to the listener was presented through sound. Two spatial displays were used in the experiment. One was a haptic pointer interface (HPI) where the user was expected to scan the environment with the use of a hand-held pointing device. If the HPI was pointed within $10°$ of the next waypoint, a fast beeping sound was heard to confirm the direction. The user could then continue in that direction with the assurance that he/she was heading in the right direction if the beeping sound was heard. The other spatial display used spatial audio filtering to convey direction. In the virtual sound condition a rapid beeping sound was spatial filtered to appear from the direction of the next waypoint.

Both displays was turned on and of with the help of a single button. When the button was pressed the sounds conveying directional information could be heard, when the button was released no directional information was given. The distance to the next waypoint was in addition read aloud of a synthetic voice every eight second.

System

The system was run on a notebook computer carried in a backpack. The prototype used a *Trimble 12-channel differential GPS receiver* that provided an accuracy of position that was below 1 meter during the tests [**?**]. This accuracy is way better than the accuracy of GPS-receivers in more portable devices, but the researcher justified their choice of a 'bulky system' with the hope that smaller systems with the same accuracy eventually will become commonplace. The whole system weighted 2.5 kilograms.

A small magnetic sensor to detect orientation (compass sensor) was positioned at the hand-held pointer to accommodate scanning when the HPI spatial display was used. The 2006 article does not specify where they positioned the compass sensor when using the virtual sound display. Earlier experiments positioned this either mounted in front of the users waist, or on top of the headset [**?**]. If the sensor is positioned on the headset all head rotation will lead to an update of the virtual sound environment. When mounted on the body a change of body orientation is required to update the virtual sound environment.

Study

The experiment had several goals. Earlier test had only been conducted on open places and not in urban areas that better reflect visually impaired pedestrian everyday routes. One route was therefore set to a urban area with intersections, sidewalks and streets. Another route was set in a more open area in a park with curved paths. The routes was intended to be traversed without aid from the researchers, where the participants navigated with the UCSB Personal Guidance System and a traditional white cane to help in avoiding obstacles.

The researchers was interested in seeing how often people would choose to use the spatial display and when they chose to do so. Would a more structural environment lead to less use of the system, than in a less structural environment as a park?

The third goal was to compare the two different spatial displays, that was designed based on earlier tests of a larger variety of displays.

Eight subjects that were legally blind was tested and a total of 30 trials was completed. The two routes was conducted with both the HPI and the virtual sound display. The time to complete each route, the amount of time with directional guidance, errors made and the distance travelled were recorded.

Results

Both spatial displays proved very successful. All participants in the study manage to follow the routes and reach the target in all trials without the need for outside assistance [**?**]. To determine how accurate they were in following the route, actual travel distance was compared to an ideal travel distance. The mean extra distance was 18%, but as low as 10% in one third of the trials. This indicate that the participants manage to follow directional information with high accuracy. One participant managed to walk past a waypoint, but discovered it and turned around to regain the path. Another participant got of course and got trapped in a bike rack. Apart from these incidences the UCSB Personal Guidance System proved to be of great use.

Street		Park	
Virtual Sound	HPI	Virtual Sound	HPI
406.7	438.5	344.3	427.5

Table 2.1: Mean times in seconds of all conditions (SD: 95.5-138.4) from **?**

The mean time of route completion for all conditions is summarized in table 2.1. The use of the virtual sound displayed proved to lead to significantly shorter route completion than did the HPI display ($t(7) = 1.80, p = .058$).

The total time the spatial displays was accessed show that the participants requested more directional guidance when in "semi-bound" environments like the park, compared to the more "bounded" environment of the street ($t(7) = 4.83, p < .001$). The amount of "travel skill" seem to affect how much directional guidance the participant request. Those participants who had the shortest completion time on each route also requested less directional guidance ($t(6) = 0.73, p < .05$).

The participants reported great satisfaction with the system. They said they wanted this types of spatial displays in commercial GPS-aids and that they would have used the directional guidance features when needed. They reported the spatial displays to feel accurate in giving directional information, that they felt it to be safe and easy to use. The virtual sound display was reported to more easy to use, but slighter less safe. The prototype gave distance information every eight second. When asked how often they wanted to hear such information the responses given were between 5 and 10 sec. The participants wanted the ability to change this frequency on demand.

2.5 Visually Impaired Pedestrians and Navigation

Navigation may be defined as the methods used by a person to find his/her way through the environment. *Visually impaired* people discussed in this context are people that have either lost their sense of sight, or suffers from a reduce sensitivity of vision that makes pedestrian navigation based on vision challenging. The loss of sight will affect the way people perceive the environment around them and give larger importance to other senses. Hearing, is together with sight, one of the senses that is best suited to perceive things from a distance [**?**]. A listener can locate vehicles like cars based on the sound they produce and detect objects and buildings based on sounds they reflect. The spatial attributes of sound normally help visually impaired people in perceiving the world around them, where sounds source get a position and size. However, ambient sounds are not always enough. Natural sounds do not tell such things as which street one is on, which brand of supermarket that is close, or when the walking path end in a steep staircase.

As with all people, visually impaired people need to get around from time to time. They need access to an environment they can comprehend and cover, and they need independence of movement where they can walk safely without the immediate help of others [**?**]. Many aids have been developed to try to fulfil these needs by compensate for the lack of sight in navigation. To understand the role of navigation aids we first have to look closer at navigation.

2.5.1 Orientation and Mobility

Two important distinctions in navigation with respect to pedestrians, is *orientation* and *mobility* (O&M). Orientation can be defined as *"knowledge of one's distance and direction relative to things observed or remembered in the surroundings and keeping track of these spatial relationships as they change during locomotion"* [**?**, p 750]. In short terms it concern knowing *where* one is with respect to the environment. Mobility deals on the other hand with the *ability of movement* through the environment and can be defined as *"the ability to travel safely, comfortably, gracefully and independently through the environment"*[**?**, E. Foulke 1997 cited in]. If the loss of sight makes it harder to avoid obstacles or walk in a normal pace, mobility is affected. Variants of the same distinction is in some places referred to as *navigation* versus *obstacle avoidance* [**?**] or *macro-navigation* versus *micro-navigation* [**?**].

> *To oversimplify this distinction, one might use a "mobility aid" to negotiate a safe path across an intersection and an "orientation aid" to determine which intersection one is about to cross.* [**?**, p 285]

Spatial Updating

A fundamental concept when it comes to orientation is *spatial updating*. The term can be defined as "the process of keeping track of the changing distances and directions to objects or places that result from self-movement" [**?**, p 46]. When a sighted pedestrian walk down a street in daylight he is continuously updated on his spatial position in relation to the environment through his vision. If he plan to take the road to the left in the next intersection, he can plan when to do his turn based on the visible distance to the intersection. When he reach the intersection and make a left turn, he may calibrate his orientation based on the orientation of the road. Imagine walking the same route without vision. How does the pedestrian know when he has reached the intersection, how does he know that he has turned enough to the left to be able to follow the next road? In the task of orientation we all need methods of spatial updating our position to be able to navigate.

Frames of Reference

When talking about orientation for pedestrians it becomes important to distinguish between two frames of reference. When a person describes a location with respect to his own orientation he is using a *egocentric* frame of reference. The phrase "Walk straight ahead and take to the right" has no meaning if one do not already know the position and orientation of the person saying so. If the person had instead described the location as "Walk to the building that is two blocks north of the church" or "Walk to the red dot on this map" he had been using a *allocentric* frame of reference. When using these forms of reference one speak of locations that is independent of one's own location.

2.5.2 Mobility Aids

The distinction between orientation and mobility comes apparent when it comes to navigation aids for visually impaired. No electronic mobility aids has managed to replace the use of canes and dogs, but electronic orientation and navigation aids are becoming increasingly popular. This section will look at the use of canes and dogs, and present three electronic navigation aids that are commercially available.

White Cane

A *white cane* (or *long cane*) is a stick that is held in hand and directed forward with the tip at ground-level. The white cane is used to detect obstacles or changes in the foundation while walking, and is normally moved back and forth in an arc in front of the user. The cane extends the arms reach with 1 to 2 meters and can be described as a powerful perceptual tool [**?**]. The cane delivers most importantly tactile feedback, but the sound of the tip against the ground may also give away cues about the terrain.

The white cane is mainly a mobility aid and has less potential in helping the user orientate in the environment beyond the reach of the cane itself.

Guide Dog

Certified *guide dogs* are dogs that have gone through extensive training to be able to guide blind owners when walking. The dog is held with a rigid curb that makes it easier for the owner to follow the dogs movements by one hand. Dogs are usually trained to follow side walks, avoid obstacles like poles or low hanging branches and locate crossings, cross walks and building entrances when instructed to do so.

As with canes, the guide dogs are mainly recognized as mobility aids. They may still aid in orientation in some forms, where they can be trained to learn a small range of destinations by name, or help the owner locate general structural elements in the environment like mail boxes and such.

Echo Localization

Some blind and visually impaired people have learned to use echo localization to detect walls, obstacles and other objects in their immediate environment. By make snapping or clicking sounds and listen to the reflections of these sounds from the environment, they are able to detect position, size and material properties of objects with high accuracy. Binaural hearing is an important part of this ability. The direction the reflections originates from indicates where this object is positioned. Spectral differences and delay may provide other information about the object that produce the reflection. This learned skill can be categorized as an advanced mobility aid. Echo localization serve as an important example of how versatile sound can be used in navigation.

2.5.3 Orientation Aids

GPS is an acronym for *global positioning system* and is a widely used system for electronic devices to determine geographical position [?]. The position is inferred through cross bearings of three or more satellites and outputs a position as degrees longitude and latitude. This position data can then be compared with an electronic map to determine the devices location with respect to other places or artefacts in the environment.

The introduction of *electronic travel aids* (*ETAs*) that utilize this form of positioning can be of great use in orientation and way finding for all people. Most systems are made for sighted users, but some devices has been designed especially for visually impaired users. Since visually impaired people methods of orientation and navigation is in some ways different than the methods used by sighted [?], it is beneficial to design devices that match the needs of this user group.

Trekker Breeze

Trekker Breeze™ is a widely used GPS Aid for visually impaired pedestrians, made by the Canadian company Humanware[27]. It is a standalone device that utilize synthesized speech to help its users orientate and navigate. The device has large tactile buttons that is easily recognizable by blind users and the device is reported to be one of the most simple GPS Aids for visually impaired that is commercial available [?].

The device can be used in two ways. Either by 'looking around' where it reads aloud street names, points of interest and intersections that the pedestrian encounters while walking. The device use locally stored electronic maps for this types of information, but it is also possible to create personal landmarks where the user records a short voice message of the point of interest. This could be everything from 'Entrance of Post office' to 'Steep staircase, use the railing on the right side', where the use is mostly up to the user.

The other way of using the Trekker Breeze is to navigate routes. A route may either be created by walking the route and record the covered trail for later use, or by letting the device calculate the shortest route to a selected destination based on stored map data. The device will then be able to guide the user to follow the route through verbal instructions.

The Trekker Breeze is mainly an orientation and navigation aid. Information that aid in mobility may be recorded as landmarks, but the device should always be used in conjunction with mobility aids like guide dogs or white cane.

[27] `http://www.humanware.com/en-europe/products/blindness/talking_gps/trekker_breeze/_details/id_101/trekker_breeze_handheld_talking_gps.html` (Accessed 17. June 2012)

Mobile Geo

Mobile Geo™ is the first GPS Aid for visually impaired that was designed as an application that could be installed on a wide range of mobile devices [28]. The application from *CodeFactory* was made for the operating system *Windows Mobile* that was used by many manufacturers of mobile devices. This operating system is now superseded by *Windows Phone* which no longer supports screen readers which Mobile Geo utilize [**?**]. The application is still a good example on a widely used mobile orientation and navigation application designed for visually impaired pedestrians.

The application has more features and can be seen as more complex than the Trekker Breeze to learn [**?**]. It lets users record routes and annotate them afterwards. Landmarks and points of interests can be created through text, which later is searchable. It is also possible to preview a route or another point of interest from another location, which makes it possible for users to get familiar with an area before they go there.

The application is interacted with through the phone keyboard, where the different keys represents different commands.

The rate and amount of information given can be controlled and route instructions can even be set to be communicated through vibrations following Morse codes. Here "-.-." represent "C" which means "Continue straight" ahead. The reported usability of this feature is not mentioned.

Sendero GPS LookAround

The *Sendero GPS LookAround*™ is a rather new application for smartphones developed by the *Sendero Group*[29]. The application can be installed on either the Apple iPhone smartphone, or smartphones running the operating system *Google Android*™.

The application can not record or find routes, but focus instead on presenting points of interest and other information about the close environment. An interesting feature of the application that is not present on the other GPS aids discussed, is the ability to get the current compass direction of the device. Instead of having to walk several meters in one direction to find your current heading based on your changes of position, the Sendero GPS LookAround use the smartphones *compass sensor* (magnetometer) to detect the current heading of the user.

A method of using the device is by shaking the device shortly to indicate that you want information. The application will then read out loud your current heading, street, city, cross street and the nearest point of interest. The types of points of interest can be changed, where the default option is restaurants.

2.5.4 Guidelines for Designing O&M Aids

? has written an article that define guidelines for designing mobility

[28] `http://www.codefactory.es/en/products.asp?id=336` (Accessed 17. June 2012)

[29] `http://senderogroup.com/support/supportiphone.htm` (Accessed 17. June 2012)

and orientation aids. Much research and development has been put into the development of such aids, but less time has been devoted to making general guidelines that can be applicable to a wide range of aids. This section will reference some of these guidelines.

> *"the device should not interfere with the primary mobility navigation aid used by a blind user, such as canes and guide dogs"* [**?**, p 385]

The development of new orientation and mobility aids should seek to complement the primary aids already in use instead of replacing them. New complementary aids should also be design to not interfere with these mobility aids. The use of canes and dogs always require one hand. If a navigation aid require two hands to be operated, it will unnecessary interrupt the users control of their primary orientation aid. This also concerns the complexity of interactions with the device. Guide dogs and especially white canes require the attention of the user while walking. New devices should therefore reduce their *cognitive footprint* and be easy and fast to attend to.

> *"what is necessary and useful in a travelling aid is to present just small amounts of information, from which blind people can improve their navigational performance."* [**?**]

Visually impaired pedestrians rely mostly on their hearing to get information about their environment beyond the reach of the cane and the dog. At the same time as with mobility aids, new devices should not interfere with this ability unnecessary. The perception of ambient sounds is very important and should be design for.

Spatial updating, where the user may get continuously feedback on their location, is beneficial in navigation, but care should be given to not overload the user with information. Only the most relevant information at the times they are needed should be aimed for.

> *"The goal is to enhance an independent navigation of blind individuals through the usage of mobile. The role that an electronic aid can play in navigation should decrease as the individual develop more M&O skills and strategies."* [**?**]

Independence of movement can be seen as common goal for all orientation and mobility aids. The O&M aids should provide enough help to let their users navigate freely in the world. It should reduce stress and anxiety related to orientation in unfamiliar environments and reduce the need for reliance of other people when it comes to navigating. The users should be given control when dealing with the aids, where they should be able to reject suggestions and interact freely with all elements in the virtual environment. When the user has developed enough skill and knowledge of navigation a particular area the role of the navigation aid should be downgraded, where it only present information that the user can not acquire in other means. This put great challenges on the design of future O&M aids.

> *"Therefore, the idea is to reuse existing devices and infrastructures for M&O purposes, ending up with a system that is cheap, available to blind people, that provides the information needed for navigation, allows them to understand the environment without the aid of a sighted person"* [**?**, p 387]

This guideline talk in favour of using traditional smartphone platforms in designing specialized electronic orientation and mobility aids. Independence of movement is here the main goal where visually impaired people should not need to rely on the help of others in navigating in their environment. Electronic O&M aids that complement other traditional aids as cane and dogs should be deviced.

Chapter 3

The Sound Guide Case

In the quest of revealing if spatial audio can be used in orientation aids for visually impaired pedestrians, a prototype was developed

> **Research Task:** Develope an application for a commercial smartphone platform that utilize spatial audio for directional guidance in route navigation for visually impaired pedestrians

The prototype developed was named the *Sound Guide* based on the use of sound as a medium for guidance. Since the prototype was based on a single concept with few development iterations it can be likened to a *case study*. The case is *explanatory*, where the single prototype is used to reveal findings that can be generalized beyond the case itself. The findings from the user tests of this prototype is later compared with theory to provide new insight in how such applications can be designed in the future.

3.1 Concept and Vision

Navigation applications for visual impaired pedestrians normally present directional information through verbal instructions. This may be sufficient in most case, but there are many situations were a spoken directional cue, like 'take the first road to your left', is ambiguous or too simplified. How do you determine when there is a road to your left if you can't see? Is left always the same as a 90° turn? What if an intersection has more than four exit roads?

A method of giving directional information that provide richer and more intuitive guidance is needed. Spatial audio could provide such a method. Instead of following instructions given through verbal clues the users could be instructed to follow the sound itself.

This prototype was developed to be used by visual impaired users and was intended for navigating prerecorded routes. A route is a walking path from A to B that consist of several waypoints (decision points). The waypoints is positioned at intersections or other places where the route changes direction. The application use its sensors and localization-capabilities to determine its own location and the heading to the next waypoint. This heading is presented to the user as an audio icon that seems to originate from the waypoint itself. The spatial placement of the audio icon is provided by HRTF filtering and use of headphones. This audio icon will be referred to as the *audio beacon* from now on - since lighted beacons used at sea is a good analogue for directional navigation by the help of stationary objects.

The users holds the device in one hand while they walk ,and get route directions as an augmented audio beacon positioned in the environment. The application only deals with road navigation, and therefore needs to be combined with a guide dog or a walking stick to help users avoid obstacles or dangerous areas. It should therefore be rated as an orientation aid and not a mobility aid. The prototype is developed to be used in testing situations where initial setup and configurations is done by a facilitator and not by the visual impaired user. The limitations of the prototype also made it necessary that the users were assisted in some aspects of the navigation tasks.

3.2 Software Developments

The prototype was developed as a standalone application on a commercial available cellphone model. The prototype was developed for the sake of this study in a development period that lasted about two months. An agile, but informal method of system development was adopted. Since there existed no documentation of similar applications for iPhone, the development followed an exploratory approach where frameworks and designs were tested and then rejected or adopted during development. The evaluations was mainly *opportunistic* [**?**], where friends, colleges and supervisors was asked to comment on ideas or challenges during development.

The specifications that guided the development of the prototype was not fixed, but changed dynamically during development when relevant theory and technological limitations provided new challenges. Figure 3.1 show four approaches to the design and development of the Sound Guide prototype. Technological support (1), limitations of perception (2) and design principles (3) all guided the development and the resulting design of information presentation in the prototype. The users goals and needs (4) was first fully approached later in the process and served as the main basis for evaluation.

3.2.1 Technologies Used

The resulting frameworks and technologies used should only be seen as one possible approach. Other possible methods of implementations is discussed in section (see 6.2.5).

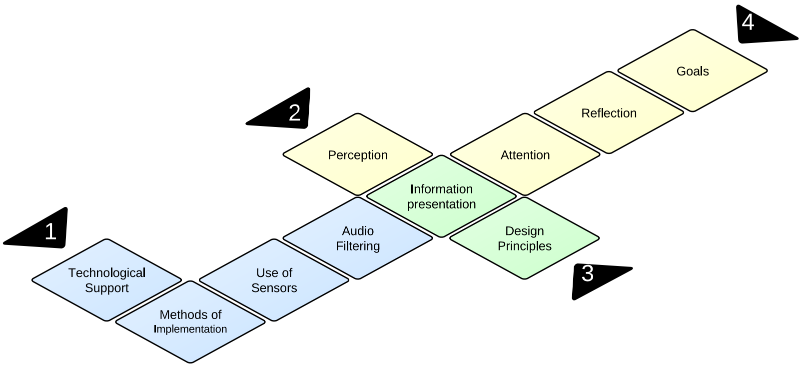

Figure 3.1: Illustration of the multiple theoretical approaches to the design and development of this prototype and the intermediate steps.

Apple iPhone 4S

The smartphone Apple iPhone 4S™ with its operating system iOS 5.1, was chosen as the device platform for this prototype. This device has a large market share of smartphones [**?**] and should be able to represent the technological possibilities on commercial available devices today. The iPhone have a dual-core processor on 1 GHz, wireless network connection, GPS localization and sensors like gyroscope and compass sensor, which is used in this prototype [1]. App developers for the iOS platform get access to lower level libraries which gives control of sound playback with low latency. Applications developed for the iOS platform can be distributed to all iPhone owners through the App Store™ marketplace. This makes it possible to distribute specialized assistive applications with a global reach at a low cost.

Apple iPhone earphones

For sound playback the earphones included with the Apple iPhone 4S was used (version MB770G/B). This earphones does not deliver the best audio quality, but was chosen to represent the possibilities of standard earphones on the market. Another argument for the earphones was the low level of isolation they give from external sounds. Ambient sounds from the environment are practically not affected when the earphones are in place. This is an important point for visual impaired users that rely on their hearing to detect obstacles and dangers, like moving cars and people, in their environment. The choice of earphones will be discussed further in section 6.1.7 and 6.2.5.

[1] `http://www.gsmarena.com/apple_iphone_4s-4212.php` (accessed 18. May 2012)

FMOD

To gain better control of the design and playback of audio, the proprietary audio library *FMOD™* was used. FMOD is widely used in the gaming industry for delivery of interactive audio on multiple platforms[2]. Their *FMOD Ex Programmers API* (version 4.38.02) was used to control the audio output in the prototype, while the accompanying program *FMOD Designer* (version 4.38.05) was used for sound design[3]. The interactive nature of the sound library and the flexible sound design editor, suited the project well.

Panorama 5

The spatial audio filtering was done through the use of the audio plugin *Panorama 5™* (version 5.60)[4] from *Wave Arts Inc*. This plugin was used through an audio editing software running on a computer to make sets of audio files with different spatial location. Panorama 5 is an advanced digital signal processing tool that can simulate binaural cues together with acoustical attributes like room reflections, reverberation and distance attenuation. The HRTF set that was used in this prototype is based on recordings from a dummy head called *Knowles Electronic Manikin for Acoustic Research (KEMAR)* that was measured at the MIT Media Laboratory and included in the plugin package [**?**].

3.2.2 Implementation

The prototype was written in *Objective-C*, the native language for application development on the Apple iOS platform, and compiled using Apple's own integrated development environment *XCode* (version 4.3.2). The resulting prototype application was installed on a iPhone side by side other native applications (also called *apps*). This source code of this application is made available for the community (see the appendix).

Direction and Navigation

One view of the prototype was made to record routes by a manual process. The route is recorded by walking the actual route and storing the GPS-coordinates of waypoints in optional key positions. The current GPS-location of the device is fetched from the *CLLocationManager* class in the *CoreLocation.framework* that is a part of the native *Cocoa Touch* development environment for iOS applications. This made it possible to see the current position of the user relative to the stored waypoints of the route. A map view was implemented to display the position of the waypoint as a method of validating the route. The map had no other important functions and was turned of during testing.

To give correct heading instructions, the current heading of the user had to be determined. To accomplish this a combination of several sensors

[2] `http://www.fmod.org/fmod-aboutus.html` (accessed May 18. 2012)
[3] `http://www.fmod.org/fmod-downloads.html` (accessed January 12. 2012)
[4] `http://wavearts.com/products/plugins/panorama/` (accessed January 12. 2012)

was used. The heading of the device was taken to reflect the heading of the user, as if the user was holding the device in front of him/her with the screen facing up. The compass sensor in the iPhone was accessed through the same class as the GPS-coordinates (*CLLocationManager*). This sensor registers magnetic fields in the environment and can detect the current compass-heading of the device. Early experiments in the development process revealed that the heading information from the compass sensor was slow to react to active rotation of the device, and sometimes appeared to jump to misleading headings that were off by up to 180 degrees.

To counter the weaknesses of the compass sensor, the gyroscope was introduced. The gyroscope is accessed in the *CMMotionManager* class in the *CoreMotion.framework* and can provide the device attitude relative to the world. When the device is actively rotated, the rotation information from the gyroscope is much faster and responsive than the compass sensor heading. The gyroscope doesn't detect the compass heading directly, but is precise in detecting changes in rotation over short timespan.

The gyroscope will start to drift over time. A short walk up and down a hallway at the University could produce drifts ranging from 1 to 20 degrees off. A combination of the sensors were therefore used to best detect the current heading of the user. The gyroscope was used as a base, where its heading was updated by the compass sensor when this had been stable for some time. This made the heading information fast and responsive, but misleading effects from the sensors still kicked inn occasionally.

The GPS position of the user was used to detect the compass-heading to the next stored waypoint. When this was subtracted from the users current heading, a heading to the next waypoint relative to the attitude of the user was achieved. For testing purpose a graphical arrow was made in the interface to display the direction to the next waypoint. When the user appeared to be closer than 1 meters of the next waypoint, or had walked by it, the direction to the next waypoint in line was presented. In this way a turn-by-turn navigation aid could be made.

Spatial Sound Interaction

The concept of the prototype implies that the heading information described above should be displayed through an audio beacon with a corresponding spatial position. To achieve this different methods of implementation could have been made, as will be the subject of later discussion. A preprocessed audio approach was chosen for this prototype.

An audio icon was selected to be the communicator of the spatial position. The audio file, that was in high quality in a lossless format, was processed through the Panorama 5 plugin described above to originate from one specific location on the horizontal plane around the user. There were made a set of 15 high quality audio files with the *AIF* file format based on the same sound clip. These were through Panorama 5 distributed in spatial positions 24 degrees apart to form a complete circle around the listener. Each audio file would then represented a section of this circle.

The prototype should be able to load and play the audio file with a spatial location that best describes the heading of the next waypoint. Since the filtered audio icon has a length of half a second, it needed to be able to change spatial position while playing to accommodate the user turning around. If the audio was locked to one position related to the user during rotation, it seemed to break the illusion of a sound event located in the environment and localization appeared worse.

To allow a seamless skipping between audio files while playing, the audio library FMOD was used. The set of audio files was imported into FMOD Designer and placed on top of each other in a multi-track event. A screenshot of this can be seen in the appendix. The horizontal axis of the sound event represent a heading from -180 to 180 degrees. Each file have a base volume level of 0, that increase to 100 at the heading it represent. To deliver a level volume around the user the audio files was faded linear into each other. This multi-track sound event was then exported to be used with the FMOD Ex Programmers API for iPhone. The FMOD API delivers simple methods for interaction with sound events through Objective-C. The event could be played and the direction, with the accommodating audio files, could be changed during playback without noticeable glitches.

When this setup was tested with headphones and the iPhone held in front, the audio beacon appeared to sound from a location in the environment and keep that location even during rotation and movement. This method of implementation managed to deliver a simple, but believable interactive augmented reality based on sound. In the tests conducted, the users had to hold the device in front of them with one hand. Ideally it should be possible to have both arms free by carrying the device in a strap around the neck, but due to technological challenges this was not possible with the prototype. This will be discussed further in 6.2.5.

3.2.3 User Interface Design

In line with the research questions, only the part of the prototype that utilize spatial audio will be discussed in this section. This was also the only part of the prototype that was tested on users.

The Audio Beacon

The concept implies that users should follow the spatial location of a sound when navigating outside. What type of sound should that be? We posted a similar question on social networks and got a range of different suggestions from friends and family. Some respondents wanted sounds that was soft and pleasant, other wanted sharp and distinct sounds. The examples ranged from clicks, running water, bird songs, laughter, voices, wind, summing of bees, bubbles or music. These feedbacks was mostly used as inspiration.

Due to the preprocessed approach the audio beacon had to be short and repeatable. Several sounds were tested based on how well they communicated spatial position and how pleasant they was to listen to. A sound icon used by Apple's operating system OS X Lion was selected. It is called *Purr.aiff* and is found in *System/Library/Sounds* and sounds like a mix between a sonar beep and a cricket call . It is distinct and soft in tone and we found it pleasant to listen to even on repeat. A frequency spectrum analysis of the audio file reveal a broad spectrum with intensity across the frequency range audible for human beings (about 20 Hz to 20 kHz). As mentioned in section 2.2.1 spatial audio localization benefits from sounds with broad frequency spectrum. This will be discussed further in section 6.1.2.

Panorama 5 filter settings	
Mode:	headphones
HRTF:	MIT Kemar
Coordinate system:	Polar
Elevation:	0°
Azimuth:	0-360°
Distance:	20m
Reflection:	disabled
Reverb:	disabled
Spatial mix:	default

Table 3.1: How the audio beacon was filtered in the prototype through the Panorama 5 plugin from Wave Arts Inc.

The audio beacon was given a spatial dimension through the Panorama 5 plugin. Effects like reflections and reverberation was not used and only binaural and spectral cues was applied. This type of spatial sounds is called *free-field sounds* and may be likened to sounds perceived from mountain tops, snow clad fields or anechoic chambers where wall and ground reflections is not present [**?**].

If reflections is used, the length of the delay affect different aspects of the listening experience. Short reflection delays up to 50 ms increase the perceived depth and width of the source, while reflections above 150 ms create the impression of spaciousness in the environment [**?**]. Since the goal of the prototype was to direct users with the greatest possible accuracy in horizontal localization, reflection- and reverberation-effects was not included in the audio filtering.

While the elevation of the audio beacon was stable at head level, the azimuth (angle of heading) was changed according to the intended direction it should illustrate. The motion sensors of the device detected rotations initiated by the user, and updated the audio playback with a similar change in spatial position of the virtual sound source.

This process will naturally cause a small latency in the user rotation and the accommodating movement of the sound source. This appeared not to be problematic, where the position of the virtual sound source appeared stable. This is confirmed by **?** which report that latency on the position update of sound positions based on head movements less than 50 ms is not detected by listeners. The actual latency of the prototype was never determined.

The audio beacon was repeated every other second to make navigation more predictable and provide spatial updating while walking. Just as with lighted beacons used at sea. The use of a repetition is favourable on sound localization where subjects evaluate the position of the sound sources with greater accuracy if they are familiar with the test signal [**?**]. Sounds with a consistent character, like voices, musical chords or running water, may also benefit from this.

Initial test revealed that it was difficult for users to determine if the audio beacon was in front of them, or to the back of them. A lowpass filter was put on the sounds from the rear to make it more easy to distinguish sound from the front compared to the rear. The lowpass filter started from 108° offset from the center and grew stronger in the back. This enhancement of spatial cues will be discussed further in section 6.1.5. This can also been seen as a *filtear* effect, where sound filtering is used to mark sounds with attributes [**?**].

User Interaction

The initial concept for the audio beacon was to appear from a fixed length in front of the user, and only communicate the direction towards the next waypoint. This could be likened to following after a person in the dark just by the sound of his/her voice. A turn in the road is then only apparent when the audio beacon suddenly change direction, and can not by anticipated in advanced based on the sounds given. What if the audio beacon was fixed to the position of the waypoints themselves? Could the distance to the next waypoint then be communicated by the intensity of the sound? In the later phase of development the volume level of the audio beacon was changed according to the distance. When the next waypoint was far away the audio beacon appeared on a fixed low volume. When the distance went below 20 meters the volume slowly grew stronger as the distance decreased. The use of reflections and high-pass filters could have improved the localization of distance according to the literature (see section 2.2.1), but a simple change of volume was regarded as a sufficient indicator in this setting.

When people walk outside they can normally look ahead and see which direction the road turn before they reach the turning point itself. This makes the trip more predictable and perhaps makes the route more easy to recognize and follow. Could this form of information be presented with the concept of audio beacons? A mode was made that in addition to the main audio beacon at the position of the first waypoint, also presented the location of the second waypoint in line. The second waypoint ahead was presented with the same audio icon, but it was lower in volume than the first and appeared 750 milliseconds after. This made it possible to detect how the road turned, and get an impression of the length of the next section.

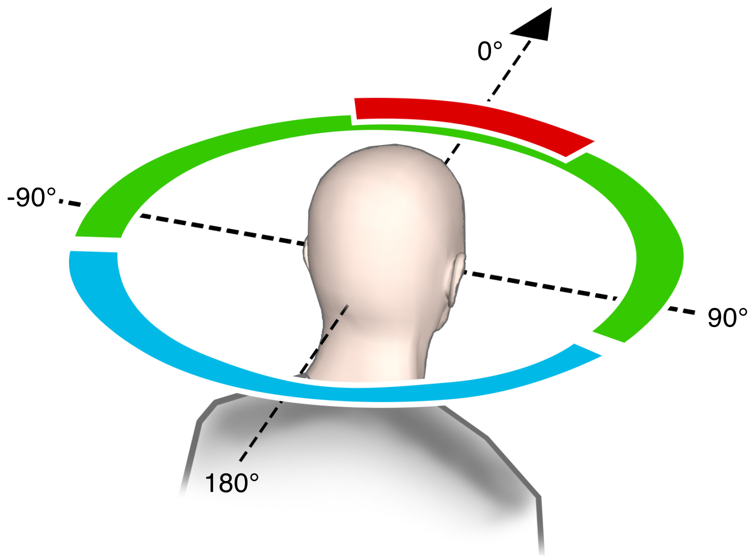

Figure 3.2: The sound design of the prototype visualized. The *green* section goes from $-108°$ across $0°$ to $108°$ and represent the audio beacon with spatial filtering. The *blue* section in the rear is where the audio beacon is applied a lowpass filter to reduce front/back confusion. The *red* section represent the heading confirmation that goes from $-20°$ to $20°$ and double the audio beacon.

Heading Confirmation

After the initial pilot tests of the prototype an important addition to the audio beacon was added. Users reported uncertainty in determining if the audio beacon was straight ahead and that they were at the correct course. After suggestions from my external supervisor at Media Lt, a second audio icon was added to provide better feedback. When users turned towards the direction of the next waypoint, a click sound was introduced. The click sound was a recording of fingers snapping. It was not processed as spatial audio, but was only a balanced stereo track. Together with the audio beacon the click sound was repeated when direction offset was below 20 degrees. This direction feedback will be discussed further in section 6.2.3.

Route Completion

When the route was completed and the users had reached the last waypoint, another audio icon was introduced as feedback on route completion. All other sounds ended and a simple three note melody was played. The melody is used in some systems to indicate success or that an event has finished. This sound ended the part of the prototype that was tested on users.

3.3 Comparisons with other Applications

The widely used *Trekker Breeze* aid, presented earlier, is a standalone device especially made for the purpose of GPS navigation. This makes it possible to rely on tactile buttons with dedicated functioning that should be easy to learn and recognize for visually impaired users. The Sound Guide is developed for an open smartphone platform, where all tactile buttons is mainly used by the operating system. The main method of interaction with the Sound Guide prototype is through the touch sensitive screen and its Voice Over functionality.

The Trekker Breeze have a wider range of functionalities than the Sound Guide prototype, but does not provide real-time directional information to the user. This is also true for the *Mobile Geo* aid. The user may be told which compass direction that have been traversed the last ten meters, but the responsibility of keeping a straight course is very much up to the user. The Sound Guide prototype deliver continous directional information in high detail relative to the next waypoint.

As with the Sound Guide prototype, the application *Sendero GPS LookAround* is made for open smartphone platforms. The application can not record and navigate routes, but is made to present information about the close environment of the user as street names and nearest point of interest. The application also use the compass sensor to present the current heading of the user. This heading is only with an allocentric frame of reference (north, south, west etc) and not an egosentric frame of reference (the waypoint is to your right) as with the Sound Guide Prototype.

The *UCSB Personal Guidance System* prototype (described earlier in section 2.4.2) share many similarities with the Sound Guide prototype on the conceptual level. Both prototypes use spatial audio icons to convey directional information in route guidance tasks. The routes consisted of several waypoints positioned at each intersection in both prototypes. When a waypoint is considered reached the next waypoint would be presented.

The Sound Guide prototype could present the location of two waypoints ahead, to provide information of the coming turns. This functionality was not present in the UCSB prototype. Distance to the next waypoint was given as verbal cues every eight second in the UCSB prototype, while this information was only presented as intensity differences of the audio icon in the Sound Guide prototype. The users were also given the functionality of turning on and off guidance in the UCSB prototype, while this functionality was not given in the Sound Guide prototype.

When it comes to implementation there is large differences on these two prototypes. The UCSB Personal Guidance System was developed for a laptop that was carried in a backpack, connected to a high precision GPS receiver and an external compass sensor. This setup provided much more precise position information than what we accomplished through our use of traditional smartphones, but can be categorized as bulky, heavy and less practical for everyday use.

The Sound Guide prototype can be categorized as an highly unique route gudiance aid with respect to the use of spatial audio, orientation sensors and smartphone as implementation platform. The UCSB Personal Guidance System is the closest match we have found to the conceptual design of our prototype. A more thourough investigation of the different aspects of the Sound Guide prototype will be done in the discussion.

3.4 Use Scenario

To summarize how this application concept was designed to be used, a small scenario will be described. Imagine a visual impaired woman that want to walk from her home to the nearest supermarket. She has recently moved to the area and is not that familiar with her neighbourhood. Her sister have helped her record a route to the supermarket by walking the route and dropping waypoints at each turning point. Now the woman want to follow the same route with the help of the navigation application on her smartphone. She select which route she want to follow, hangs the phone around her neck and puts on headphones. She brings her walking stick and go outside. The following functionalities of the prototype are covered in following users tests and discussions.

When standing outside her house she can hear a sound that appears to come from her left hand side. She recognize it as the applications audio beacon and intuitively detect the direction of the sound. She turns toward the sound until it originates from her front and the sound icon changes to a clicking sound. She start to walk in that direction and use the walking stick to avoid obstacles. The clicking sound continues to repeat every other second and tells her that she is still on the right course. After a while, she starts to fall off course. The clicking sound then disappears and the audio beacon appear slightly to the right. The woman correct her course accordingly. When a change of road is soon to come she can hear that the waypoint is coming closer. At the same time she hear a second audio beacon originate from her left, which tells her that she shall soon change direction and follow a road to her left. When she reach the intersection the location of the second beacon becomes the first and she can hear the next waypoint in line in the far distance. After a couple of more roads is covered, she reach the location of the last waypoint and a small melody sounds. She then knows that the route is covered and she has to stand outside the supermarket. She removes the headphones and finds the entrance of the supermarket by following the natural sounds of rolling shopping carts, chatting cashiers and sliding doors.

Chapter 4

Methods

This chapter outlines the different methods and techniques used to answer the research questions in this thesis. Since time and resources were limited a small number of methods was selected. Each method was selected to provide most relevant data possible to address each question.

4.1 What is a Method

Knowledge is gathered through the use of methods. Scientific methods are methods that are made to produce reliable knowledge through empirical observation. Which methods to use depends on the research questions that should be answered. The methods may be divided into two broad categories: quantitative and qualitative.

4.1.1 Quantitative Methods

Quantitative methods seek to find results based on numerical data. The hypothesis is numerical, the observations are made numerical and the results are numerical.

Quantitative methods is either *experimental* or *nonexperimental*. Studies using questionnaires or archival records are non-experimental. By looking at correlations between variables, relationships may be identified: changes in A accompany changes of B. This types of studies is very common, but they have no way of determine clear cause and effects relationships [**?**]. We may not know if A influence B, or B influence A, or if A and B is both influenced by another variable not accounted for. To gain knowledge of causal relationships we have to use experimental methods.

In experimental research a causal relationship is only established if the measured difference between conditions are *statistical significance*. Even if the dependent variable reach a higher mean in one condition than in another, there is always a chance that this difference is a product of normal variance in the measurements, and not caused by any change in the independent variable.

Scientific methods deals with this form of uncertainty in a number of ways. If the number of results is large enough the average of results from the *sample* should be close to the expected average in the *population* [**?**]. A large number of tests is therefore advised when doing quantitative research. By using statistical calculations, that take the number of results and its distribution into account, it is possible to predict the possibility that a difference in average is caused by something else than chance. There is many thresholds used, but a result is often coined as statistical significant if there is less than 5% chance that the difference in sample average is caused by chance alone.

4.1.2 Qualitative Methods

There exist research questions where quantitative methods fall short. Not all real life phenomena may be replicated inside a laboratory or be captured in questionnaires. With a *qualitative* approach, less focus is made on counting, rigid measurements and calculation. Through broad observation and qualitative analysis, themes, trends, and relationships may surface that earlier was not expected. This is especial relevant when looking at interactions between technology and ways of life, where it is hard to abstract the use of technology from the context it is used in [**?**].

> *"With qualitative research, the emphasis is not on measuring and producing numbers but instead on understanding the qualities of a particular technology and how people use it in their lives, how they think about it and how they feel about it."* [**?**, p 138]

The qualitative methods used in this work is interviews and observations. The empirical data from these methods consists of written observations, audio recordings and pictures. Statistical tests is unfit for these types of empiri and a researcher has to rely on less rigid techniques for content analysis. Content analysis is about making sense of the recorded material by categorization and coding of elements [**?**]. Such analysis give purely descriptive results and can not test causal relationships in the same way as experimental research [**?**]. The aim is often to make sense of other people ways and lives and *"work 'with' rather than 'on' people"* [**?**, p 28].

Qualitative methods are sometimes criticised for their reliance on what appear to be 'subjective' conclusions [**??**]. Less rigid methods give more responsibility to the researcher. According to **?** a researcher has to defend the validity of their study by showing that it is theoretical sampled, saturated and adequate: The subjects approached should be chosen on the quality and relevance of the information they can provide, the study is theoretical saturated when an increase in subjects approached do not seem to add new accounts to the topics, and the study is theoretical adequate in the sense that other relevant research is evaluated to give multiple perspectives on the topic.

4.1.3 Terms and Constructs

To help in the discussion of methods used more terms related to research designs needs to be introduced. Important attributes that needs to be considered in the use of scientific methods is their internal and external validity [?]. The *internal validity* of the method is the ability it has to test the hypothesis or research question it was chosen to test. If the results of an experiment is influenced by unrelated variables other than the independent variables, the experiment suffer from low internal validity and should be redesigned and redone. *External validity* relate to the degree the results of an experiment or a study can be applied to describe the real world. Is the result generalizable or was the research setting too artificial to produce relevant behaviour or experiences? Naturalistic observations in real settings have high external validity in that it portrait the real world directly [?], where results from highly constrained laboratory experiments may have less relevance in describing real live phenomena.

Triangulation is used when you validate your findings by using multiple methods to study the same topic. If results are consistent from several methods it also becomes more valid and reliable. This is also the case when you find references of other research that may support your results [?].

4.2 Methods

An *interpretive* approach has been adopted for this study. An understanding of how the technology may influence the users and how the users may influence the technology, is gain through language and the meaning people assign to the topic [??]. Through the use of qualitative methods peoples reactions, hopes and interpretations about a technology and its context can be gathered and analysed. The aim of these studies is to understand the context of the information system and in which ways the informations system influence and is influenced by that context.

4.2.1 Semi-structured Interviews

To get background information on current use, needs and experiences related to navigation for visual impaired users, interviews was chosen as method. As described in ?, p 21 *"researchers make use of interviews when they wish to obtain more detailed and thorough information on a topic than might be gleaned from a questionnaire."*.

The interview was *semi-structured* in the way that the interview guide only contained a list of broad topics to discuss [?]. This loose structure can help the interviewed feel at ease and let topics emerge naturally [?]. This approach enabled the surfacing of new topics and concerns that probably would not have been spotted using more formal methods. That makes it a qualitative method where the net is thrown out wide and more control is given to the interviewed. To accommodate this, the interview followed a *student-tutor approach* where the interviewed were treated as the expert which the interviewer wanted to learn from [?].

4.2.2 Field Studies and Observation

To get better understanding of how new technology can influence a complex activity as pedestrian navigation, field studies in natural environments was conducted. **?** describe four goals of field studies: identify new areas for technology, identify design requirements, reveal how to introduce the technology in the context, and evaluate the technology. All these goals are of interest in this study, which deals with the use, introduction and design of new technology. The late introduction of users in the development period still makes evaluation the main approach to the field tests conducted. By evaluating how a technology is used, thought of and integrated in its natural setting, a richer account of the possibilities of the technology in real world use can be studied. Field studies is great starting point for evaluation of new technology, and to help find issues that can be tested more thorough in more controlled settings, like in a laboratory.

Field studies deals with qualitative methods. The study is conducted in a natural environment of the activity and have an open approach with few restrictions. With respect to information research it is often peoples use of technology that is beeing studied. Observation is used to describe how people use the technology and how they react to different parts of the design. Audio or video can be recorded and notes can be taken. Observations is conducted either by passive observation without interference, or by participant observation where the researcher is more engaged in the activity and dialogue.

4.2.3 Content Analysis

The interviews and the field studies was tape recorded and transcribed. Written observations from the field studies was combined with the transcribes and this formed the basis for the qualitative data. The goal of content analysis is to take a closer look at the data to discover new themes and patterns that is not apparent at first [**?**]. Open coding was used as a method to label sections and utterances in the material to find common tendencies and attitudes across subjects. This was done using coloured pens on printed records. It is recommended to let the labels and categories be 'found' in the material and not imposed a predefined set of categories [**?**]. More selective coding was also used to mark sections based on relevance to different aspects of the research questions. Comments and annotations was added on the same papers in an iterative process to develop an understanding how the material may be interpreted. A *hermeneutic* mode of analysis was adopted where an understanding of the material was accomplished through an interpretation of the meaning of its parts as a way of making sense of the whole [**?**]. The codes and annotations on the parts was therefore used to understand the general tendencies and relations in the experience and opinions. Since the numbers of subjects was limited the findings was presented more as opinions and indicators than as general tendencies.

4.3 Ethical Considerations

In studies where voluntary participants are involved, several ethical considerations need to be addressed. **?** summarize three important elements of ethical considerations: participants vulnerability, informed consent, and privacy, confidentiality and trust. Care should be taken to ensure that participants understand that participation is completely voluntary and that they have the right to refuse to participate if aspects of the study makes them uncomfortable. Informed consent concerns giving information and getting participants approval. All participants should be informed of the purpose of the study and how their contributions will be treated. This is normally presented on a paper that is signed by the participants to indicate that they understand what is expected of them and understand their rights. The third element concerns the protection of personal information. All data recorded of participants should be stored according to legislations, and destroyed or de-identified when the study is ended. This should be done to ensure that individuals in the study can not be identified and sensitive information be leaked.

This study was reported to the Norwegian Social Science Data Services (NSD) [1], which is the Data Protection Official for Research for Norwegian research institutes. NSD ensure that the necessary considerations are taken into account when it comes to collection of personal data for the subject of research. Since the aspect of visual disabilities of subjects are considered to be sensitive personal information, this study was subjected to notification.

4.4 Research Design

The aim of this study was multiple. First we wanted to get a broader understanding of the context we were designing for. What methods do visually impaired people use when navigating outside, what are their challenges and what are their needs? This was gathered through a semi-structured interview. Secondly we wanted to see how the Sound Guide prototype developed would fit into the context of pedestrian navigation. Did the features of the prototype meet the needs of the users? How did the participant adopt the prototype and what were their experiences? This was done as a field test where the participants were asked to use the prototype on a short route following a footpath. To control the participants perception of the spatial audio used, a preliminary test was conducted to see if they manage to detect directions with sufficient accuracy.

4.4.1 Participants

The participants were contacted by the company *Media Lt* through their existing networks of visually impaired test users. Media Lt[2] is a company that conduct research, consultancy and training on the field of information

[1] `http://www.nsd.uib.no/nsd/english`
[2] `http://www.medialt.no`

systems and the disabled in Norway. They had a supervising role in the early stages of this study and helped in planning user involvment, finding possible participants and provide a meeting room that was used during the interviews.

> *"A sighted user can not simulate the behavior of a blind user because their cognition and mental models are not alike"* [**?**, p 375]

Three female adult participants participated in the study. They were all blind, were they needed mobility aids like canes or dogs to navigate outside. They all participated in conjunction with another usability study held at the same location. No payment or other forms of reward was given for participation.

The participants were informed of the purpose of the study and their rights as voluntary participants in a research study. A consent form was signed after been given all information verbally.

4.4.2 Time, Location and Logging

The study was conducted in a meeting room at Media Lt, where the field test was conducted outside in a neighbouring residential area. It was held in the spring, in warm weather at daytime. The study lasted for approximately one hour and was facilitated by one researcher. One participant participated per session. An audio recorder was started after the consent form was signed, and provided audio logs for the whole session. Observation notes were written during observations and after the session.

4.4.3 Test of Sound Localization Accuracy

To see if the spatial audio used in the prototype could give sufficient directional information for route guidance, a test was first conducted. The participant was asked to stand in the middle of the room wearing headphones, while holding the iPhone flat in level with their waist. The participant was first presented with a playback of a simple stereo sound clip and asked if he/she heard the sound with equal strength at both ears. This was done to reveal if the participant had any asymmetrical hearing loss that could affect binaural hearing.

The participant was then instructed that he/she soon would hear an audio clip through the headphones that would repeat every 1,5 second. The audio clip was made to sound coming from outside the head in a random direction. The participant was asked to identify the horizontal position the sound originated from. This was indicated by turning of the body, and the smartphone that measured the heading, in the direction of the sound event. The direction was confirmed by either pushing a button on the touch screen or by giving the device a small shake. The perceived direction of the sound event is in this test operationalized as the direction the participant indicate the sound event to originate from by the heading of their body.

This localization task was given thirteen times in a continuous session. When finished the participants were asked how they experienced the task. A second session was administered where the spatial audio clip was only played once for each localization task. This made the task harder, since continuous feedback of sound position while turning no longer could be used. The sound icon used in this test was the same sound as the one used for the sound beacon in the Sound Guide prototype.

These two tests were also conducted with an audio clip that was not filtered with HRTF, but only produced a simple stereo panning. The purpose of this was initially to compare filtered audio against panned audio to indicate the need for HRTF filtering in such applications. This was later found to fall outside the aim of this study, and no focus will be given to it later in this thesis.

An application was developed for the administration of this test. It was based on the core features of the Sound Guide prototype where turning of the device gave updated spatial positioning of a sound icon positioned at a random point in an imaginary circle around the listener. When participants indicated a direction, the angular offsets from the target was recorded by the application for later analysis. This made it possible to determine the mean accuracy of the perceived angular direction in the horizontal plane. This source code of this application is made available for the community (see the appendix).

4.4.4 Interview

After the test was completed the participants were asked to sit down for an interview. The participant was asked questions concerning how they experienced the spatial audio in the earlier test. They were asked about the use of cellphones and headphones in their daily life. And mainly they were asked to talk about how they navigate outside, which methods they use and their experience with using technology aids to aid in orientation and mobility.

4.4.5 Field Test

The last half hour of the session was spent outside. The participants were asked to bring either a white cane or a guide dog to aid them in the task. They were first led a hundred meters to the start point of the route that was prepared for the test (see figure 4.1). The area was described for the participants as a residential area with footpaths between houses. The route avoided streets with car traffic and contained no stairs or other structures the participants could fall down from. The foundation of the footpath was asphalt surrounded by grass or by picket fencing. The route was a couple of hundred meters long with a few crossings and a few twist and turns. Further directions was not given.

The goal of the task was to see how the Sound Guide prototype managed to aid the participants in navigating the route between the houses. The participants were asked to apply the headphones and carry the

Figure 4.1: The route the participants walked during testing. It was intended to be traversed from one to seven and back again. The short timespan of the test made us to start from waypoint three for the last two participants. Satellite photo from the Norwegian Mapping Authority (NMA) web service http://www.norgeskart.no/.

iPhone in one hand of their choice. The other hand was used for operating the cane or the dog. To let the participants get familiar with the auditory display used, they were asked to slowly turn around their own axis to get an impression of how the direction beacon sounded from different angles. During this task the researcher positioned himself in the direction of the virtual audio beacon to give an additional reference point. The participants were then explained how the route consisted of several waypoints. When the position of a waypoint was reached the prototype would guide them to the next.

The participants were instructed to express their thoughts, interpretations and experiences, while doing the task. They could ask for assistance any time they wanted. The researcher followed the participants throughout the route, observed the participants' behaviour with the prototype, asked questions when interesting behaviour emerged and aided the participants in overcoming the technical problems of the prototype. Audio was recorded throughout the field test and analysed together with the interview.

The target of the route was an open area surrounded by shops. When the target was reached the participants were asked how they had experienced the task. The same route was then used again in reversed direction. When the original starting point was reached the session was ended and the participants led back to Media Lt.

Chapter 5

Results

This chapter will present the results from the research methods used. This includes the test of accuracy of spatial audio perception, the semi-structured interview conducted and the field tests with the Sound Guide prototype. The results from the interview and the field test are analysed with content analysis. The following categorization of topics and their relations are based on this analysis and does not necessary reflect the theory on the topic. The results stated in this chapter will be discussed in the following chapter with respect to the research questions stated and the relevant theory.

5.1 Test of Sound Localization Accuracy

The application developed to test the accuracy of spatial audio perception stored the angular offsets for later calculation. This data was then exported from the iPhone and analysed by using the statistical tools in the program *Microsoft Excel*.

One of the participants appeared stressed and uncomfortable at this test and we decided to abort it. The participant did not manage to focus at the task and struggled with keeping the headphones in place. The data from the initial tests from this participant are therefore not included with the reason of not being valid.

Of the thirteen repetitions in each group, the first three was thought as training rounds and not included. Some additional repetitions was removed when observations during testing revealed that they were not valid. This originated from touch screen clicking without the intention to do so. Four repetitions were also removed on the basis of being outliers (more than two standard deviations above the mean). These repetitions may have been the result of front/back confusion, where the participant judged the sound event to originate from the front even though it came from the back.

The administration of groups was initially randomized to balance learning effects. Since the testing of one of the participants was aborted, the resulting order of administration was HRTF Multiple, Panning Multiple, Panning Single, HRTF Multiple for both participants. This prevent us from

Sound:	HRTF Multiple	HRTF Single	Panning Multiple	Panning Single
Mean:	$15,7°$	$30,7°$	$18,2°$	$64,8°$
SD:	$13,3°$	$23,3°$	$13,5°$	$44,6°$
Sample:	17	19	19	18

Table 5.1: Results from tests conducted on two participants. *Learning effects* prevents significance testing between the groups.

doing any significance testing that compares the use of HRTF filtering against stereo panning and multiple administrations against single, since learning effects can not be ruled out.

The results from this test indicate that the perception of spatial audio position of the participants were sufficient for directional navigation. If the spatialized audio icon was repeated the two participants managed to detect its location with a mean offset of only $15,7°$ (SD: $13,3°$). If the same audio icon was only played once, without giving the possibility of head rotation or corrections of aim, the mean offset grew to $30,7°$ (SD: $23,3°$). This is still quite decent and should provide directional information that is accurate enough for providing route descriptions in pedestrian navigation. It would make it hard to walk straight on roads without the aid of canes or dogs, but probably would make the choice of roads in crossings and intersections clear enough.

It is important to point out that results from controlled tests like this conducted in a quiet meeting room, can not be directly applied to natural use environments for navigational aids. External influence like noise and stress may influence the perception of spatial position and result in less sensitivity in binaural hearing. This is one of the reasons the snapping sound was added to the auditory display of the Sound Guide prototype to indicate correct heading.

The results are also influenced by the type of headphones used and the general HRTF filtering technique used (see section 3.2). Greater sensitivity of sound localization is possible to achieve by using higher quality headphones and individually recorded HRTF profiles. A larger group of participants and larger sets of localization tasks would also be beneficial to produce more significant results on these tasks.

5.2 Interviews on Navigation

All people have the need of travelling from where they are to another location from time to time. For people that are visually impaired the loss of sight present challenges in navigation. How to know where to go, how to avoid dangers and know when you are there? Interviews was conducted to get a better understanding of the needs, conventions and context of outdoors navigation for visual impaired.

A pilot study on a visually impaired user from Media Lt was conducted the week before the main study. The result from that interview was found highly relevant and will be included here. Since this user had been involved in the project earlier, care have been taken to avoid using statements that may have been affected by this.

These results are then based on interviews with four adult participants with visual disabilities. The results gained from this interviews will be used to help explain the results from the usability tests. All of the quotes in this and the next chapter are original in Norwegian and were translated to English for this thesis.

5.2.1 Information in Navigation

To be able to navigate you are in the need of information that help you do so. Content analysis of the interviews conducted revealed the need of information in the task of navigating by foot can be categorized into two abstract concepts: position and direction. To get from A to B, it is vital that you obtain information that help you resolve both your position and the direction you are heading. One of the participants summarized it as this: *"[Y]ou should know which direction to walk in, and you should know fairly where it lies"*.

Position

When asked of what type of information the participants would like from a GPS application, their current position was mentioned more than directional information. Position is often related to street names or the distance to close landmarks. As one participant stated: *"It is always an advantage to know the name of the street, perhaps more than knowing if you walk south, east or west."*. Another participant agreed on the priority of position in what that was wanted to know from a GPS application. The participant felt indications of compass directions like north or east would not help the participant at all in navigating, but directions towards familiar landmark were found interesting.

Direction

Compass directions are one way of presenting directions. A compass direction is based on an *allocentric* frame of reference. It is a set of directions that are fixed to the environment and that require an understanding of the original position of landmarks in that environment with respect to north, east, south, west. Only one of the participants was familiar with using compass directions in navigation. The participant had a GPS that together with the current position could present which compass direction the participant was currently walking. In the participants own words:*"When you walk in a direction you will gradually figure out if you walk south, east, north or west. In some ways you have to draw yourself a map."*

Some methods of giving direction may be more directly applicable than compass bearings. If someone is asked to point out the direction of a landmark, they would point directly to it. This form of direction is relative to the current position and heading of the users on therefore based on an *egosentric* frame of reference. An understanding of landmarks position on an abstract map is not needed in the same with relative directions. If you walk in the direction that is pointed out you will eventually get there without further knowledge. As long as you know how to recognize your target off course.

5.2.2 Navigation, Orientation and Mobility

When the visual impaired participants of this study talked about outside navigation, the concern of avoiding obstacles and other dangers was often mentioned. When asked to mention information considered most important to get from a GPS application, one participant wanted to be warned if the participant unaware walked out from the sidewalk and into the street. This is reasonable feature request for such applications, but it indicates the need for a more precise definition of what we talk about when it comes to navigation aid and which role the application is stated to fulfil.

Dangers and Obstacles

When considering mobility, many obstacles and dangers was mentioned. Dangers, in this sense, are aspects in the environment that the participants considered important to avoid when navigating outside. When informed of the outside route of the test one participants asked: *"There are no curbstones to fall down from now? Stairs?"* Curbstones was also mentioned in relation to danger when they were not present. One participant had once walked out in the middle of the street because missing curbstones had not indicated a shift from side walk to traffic lane.

Traffic was often related to heightened awareness when navigating outside. *"No dangers of getting hit by cars here?"*, one participant asked concerning the route. The participants did not like to use 'noisy' interfaces when navigating close to traffic. Subway stations was also mentioned with concern. These dangers mentioned, can be superficial categorized as being hit by traffic, or falling down ledges in the environment.

5.2.3 Methods of Navigation for Visual Impaired

The participants that were interviewed used a range of different methods when navigating outside. They had to travel to unfamiliar places from time to time, and had different preferences when it came to different methods. As one participant summarize: *"I use guide dogs and a little Trekker Breeze occasionally. Yes, and a white cane. It depends on what I do and where I walk."*

Ambient Sounds

When it comes to orientation, ambient sounds in the environment were often referred to. The sound of traffic was mentioned as a way of determine where you were in relation to the street. When walking outside during the tests, several of the participants referred to route choices by the sound of other people walking in front of them. Natural sound is considered important, as stated from a participant: *"I am very focused on, as said, sounds around me when it comes to orientation"* The same participant stated to perceive properties of the surrounding space through ambient sound: *"I am used to listen and try to detect. Is there a large room around me, is it tight around me"* This method helped in identifying if the participant was located in an open space like a park, or an enclosed space like an alley of tall buildings.

> *"And then I have to orientate myself after sound, where does the traffic goes, if I arrive at an intersection with traffic lights and so on. An then I listen to see if it is open around me or if it is tight. If it is tight it is usually a house or something else around me."*

Artificial sounds in the environment was also reported to help. One participant told about a ski hotel that positioned traffic lights that beeped continuously in the ski track to help blind skiers orientate themselves. This helped the participant to orientate in the track and made it possible to ski alone. As explained by the participant: *"I felt very safe and was not afraid to get lost at all."*

Guide Dog

All participants had guide dogs and preferred using them over white canes. The purpose of guide dogs are nicely summarized by one of the participants:

> *"A guide dog is a navigator around obstacles, you have to find the way yourself. You have to know where you shall go and tell the dog 'Front', 'Right','Left' and so on. The task for the dog is to steer you away from obstacles and to walk to crossings and to find doors and such things. 'Walk to the escalator!'"*

One of the participants stated that to own a dog that knew the location of some places in the neighbourhood, where the dog guided the participant to the supermarket if it was asked to. Still, it sometimes went wrong where the dog walked to familiar places instead of following directions to new places. Another participant characterized the errors the dog did like this: *"Labrador, like I have now, is more like: It's not that serious. If there is anything in the road to the right I walk the road to the left instead, then I don't need to bother with that obstruction."* This show that the users of guide dogs needs methods of validating the route the dog guide them on to prevent such navigation errors.

White Cane

All the participant could use a white cane, but not all was comfortable with it. It could be challenging if the edges of the road was vague, like when snow on the ground made it hard to identify the foundation and made the cane get stuck in the snow. Another participant liked to use a white cane in combination with the guide dog, while walking in busy streets to avoid being walked on by other people.

> *"When you walk with a guide dog you focus on completely different things than when you walk with a cane. Then it is more the direction and the larger aspects you care about, where you shall go to the left and right and such. But, with a cane you have to follow edges and really walk in a completely different way."*

Compared to walking with a guide dog, a white cane require more of the user. When a guide dog usually will identify dangers and obstacles and lead their owners around them, the use of white canes requires that the user identify these obstacles themselves and find a way around them.

GPS

Only two of the four interviewed participants had used a GPS device before. The participants who hadn't, had been advised by others to wait until the technology was better before purchase. They where still positive to the aspect of using functioning GPS devices in navigation where they believed it could help them to navigate more freely: *"If I had such I would have walked more (..) if you get used to equipment like that and it works, it would be less frightening. You would not get lost."*

One participant that was familiar with GPS devices were early to point out their limitations: *"I absolutely use a lot of GPS. But, you have that challenge that there is road maps that lies at the base of all GPS'. You should know a little about where you walk!"* The GPS is characterized as a device that seldom is a hundred percent correct and should be used only in combination with other aids. One participant partly blame the maps that are used by GPS devices to be insufficient for pedestrian navigation:

> *"It very often falls short. So what I miss is map series that are made for pedestrians. There are a lot of geo-data that had been useful, that are not present at the maps that are now. (..) Like cattle grids, booms.. And for me it had been very useful if such things as shops and such was put in, not only those gas-stations and such as you find on the GPS today."*

Another participant that use a Trekker Breeze GPS, which is a standalone device, complain about its large size and complicated use. The participant sometimes found it embarrassing to use it in public and did not like to carry all those different devices outside: *"I don't like such 'thingies' when I am out walking, because they are a little toilsome."* The participant liked smaller devices that were easier to use and perhaps combined with other devices like the cellphone. The participant liked to use the GPS in places that were not found familiar, but the device dependent on user provided landmarks which made the use more limited.

Assistance

When learning their way around their neighbourhood the participants mention the need for assistance by friends and family that are not visually impaired. This makes it easier to insert valuable landmarks on GPS, or in learning new routes by head. The problem for some to get people to help them is mentioned as a challenge. Not all people have a large social net around them.

Work and private appointments sometimes require the participants to travel to new places they have not been before. Taxi is mentioned as one way of getting to and from such places, but GPS or simply 'winging it', is also mentioned.

5.2.4 When Navigation Fail

Sometimes things go wrong in navigation. The participants had many stories of such incidents that were found interesting with respect to the design of navigation aids:

> *I got a little lost. And then I thought: These sounds does not compare. But then, it was not much traffic so I did not have the traffic to adjust to either. Fortunately a young woman came and helped me. She said: 'You look a little confused, do you know where you are?' No, I said, that's just what I don't know.*

This participant had taken a wrong turn just days before the interview. By getting lost into a park where the participant discovered the mistake through the ambient sounds of the new surroundings. The space sounded open when it should have been close. The participant partly blamed it on the absent car traffic, which made it harder to detect the position in relation to the street. The participant had gotten help from a stranger passing by in understanding where the fault had been made and how to get back.

Other participants tell similar stories. Where the sound of traffic helps them orientate and strangers help them out. Some are not bothered by getting help from strangers, but others say that they can not rely on others helping them. A functional GPS is mentioned as a possible aid for resolving errors in navigation on their own: *"It usually goes well. But it do occurs, that the dog do a little mistake and suddenly I am off the route. Then it could be useful with something that gave me a hint like: 'Oh yes, then you have to take the turn to the right'. Pick up the route so to speak."*

5.2.5 Use of Cellphones

The participants were all active users of cellphones in their everyday life. One participant had an iPhone and was an active user of applications and had also tried navigation applications. Three of the participants had traditional cellphones with physical keys. They used their cellphones for things like calls, text messaging, calendar and time of day. Web surfing, music listening or use of other applications were not mentioned. None of them used their phone in navigation and stated that it did not contain GPS functionality, even though two of the cellphones are reported to include GPS navigation applications adapted for visually impaired users by their manufactures.

The participants who did not have smartphones were all interested in trying an iPhone. One stated concerns of how the touch screen could be operated without pushing something wrong. Other participants were sceptical because they had no one to teach them how to use it.

5.2.6 Use of Headphones

Several of the participants found earphones (headphones that are placed in-ear) to be uncomfortable. When asked how often earphones was used one participant answered: *"No, as little as possible. (...) I feel locked up by them."*. Another complain that: *"They fall out. I do not like such plugs."*. Headphones that are placed against the ears were more preferred. They were used to listen to things like radio, music, news and audio-books. They were used when training, on vacation, when waiting for the bus or in situations where they did not want to disturb the people they were with. None of them were used to wearing headphones while they were out walking.

5.3 Usability Tests of the Sound Guide

An informal usability test of the Sound Guide prototype was performed on the three participants with visual disabilities. Some results from the pilot test will also be referenced here. Audio was recorded and observations was noted, while the participants tried to navigate a route with the help of the prototype. The participants were mostly positive to the test setting and optimistic towards the prototype before use.

5.3.1 Technical Performance

The prototype was unstable during the testing and had tendencies to show wrong directions from time to time. This could be offsets slightly off the optimal direction or in the complete opposite direction. This could endure for several seconds before it once again got a stable reading and provided a correct route direction. This affected the participants use of the prototype where they often ended in the ditch or into fences because the prototype showed wrong directions.

These problems can be attributed to inaccuracy in the GPS positioning and compass sensor of the iPhone, and not the use of the auditory display developed. More about these technical challenges are discussed in section 6.2.3.

Participant - *There! No..*
Researcher - *Yes. Where do you feel the direction is now?*
Participant - *There!*
Researcher - *Yes.*
Participant - *Into the fence?*
Researcher - *It is a little quarrelsome, I do not understand why it want to go in there.*
Participant - *Perhaps it does not know where it want to go?*

Since the purpose of the test was the use of spatial audio and not directional navigation per se, we found it beneficial to assist the participants in finding the correct course when the prototype was wrong. During the half hour the test usually took the participants were assisted from 9 to 14 times.

5.3.2 Interpretations of Audio Display

The core concept of the prototype's audio display was explained before each test. Observations and questions was used to see how the participants interpreted and used the audio display.

Direction

None of the participants seemed to have problems to determine the direction of natural sounds in the environment. *"Now I hear that somebody is walking on my left hand side, I have no problems with that."* Still, to determine the direction of the audio beacon that was spatially filtered appeared not that easy. The participants soon seemed to rely on the snapping sound for direction, and paid lesser focus on the spatial sound. When the snapping sound ended due to changes in direction, it was not always that easy for the participant to determine the location of the new direction. They seemed not sure about the correct direction before they again heard the snapping sound. When asked about the perception of direction in such circumstances, some participants reported that they got a sense of direction from the audio beacon, but they were not sure until they heard the snapping sound.

The participant in the pilot study tried the prototype without a snapping sound for front confirmation and appeared quite uncertain about route directions in the start of the test: *"We shall this way, shall we not? Yes, it says so.. Shall we not?"* This became better after some training and the participant could soon walk with more certainty without the need of heading confirmation. When asked about the uncertainty on perceived sound direction, the participant answered: *"Yes, right at the start. But then, I figured that I shouldn't be that focused on details, and then it got better."*

Apart from the uncertainty, participants always seemed to locate the directions they were given by the prototype and followed its instructions fully. If the prototype had managed to give more correct directions, the participants would probably have managed to stay on the route without assistance. When the participants learned that the prototype had errors in its direction, they soon interpreted the direction with more scepticism and did not always followed its instructions fully. Even when the prototype showed correct directions into an intersection the participants believed it only pointed into the ditch and continued forward.

The prototype had a tendency to show turns too early, where it instructed participants to turn up to five meters before the intersection. The participants soon learned to anticipate this tendency, but it appeared to make it more difficult to determine when they were supposed to turn. The use of the white cane became more important to locate the actual intersection and make the correct turn. To console the developer of this error in the prototype one of the participants said: *"You can comfort yourself with the notion that the dogs do this as well."*.

The use of *Wizard of Oz* testing, where the directional information would be controlled by the researcher and not by the sensors, could be used to avoid these technical problems in user testing. A prototype that enabled this type of testing was developed after the first user tests, limited time prevented us from doing more field tests with this prototype.

Distance

To mimic the notion that the beacon sounds emitted from the waypoints themselves, the volume was changed according to the distance to the next waypoint. This feature was commented by the participant from the pilot test: *"Now I hear the target getting closer already. And then it appear stronger and stronger I see. So that's how you have been thinking."* Not all participants commented on this functionality, but two of them showed that they had understood the difference and used it to determine their location with respect to the waypoints and perhaps as a confirmation that they was at the correct course.

> Participant - *I am probably in the intersection now.*
> Researcher - *How could you tell?*
> Participant -*The sounds, because they are almost at the same volume.*

This participant was presented with two waypoints at once. The participant managed to determine when the intersection was reached based on the volume of these sounds. No information had been given in how to interpret this, where the participant reached this conclusion based on experience with the prototype. The two waypoints reach approximately the same volume when an intersection is reached and two new waypoints are presented that are far away. The changing of beacon direction would have made the intersection apparent, but it is interesting that the participant based the detection of location on the volume differences of the waypoints which endured after the direction change.

Directions to Come

Only two participant got to test the presentation of two waypoints ahead. One of the participants found it interesting and dictated the direction of future turns for each new waypoint: *"And then I shall turn right after a while."* Since the participant had walked the short route once before, some of the intersections must have been remembered. Still, the participant seemed able to read the change of direction from the second beacon feature.

The participant from the pilot test was presented with an early version of the prototype where there was no volume differences between the beacons. The participant managed to use the second beacon to determine turn directions, but did not see the purpose of it: *"I am a little uncertain of the point of the second sound. I get it that it is the next waypoint, but.."* More test should be made to identify if this is an actual need of visually impaired pedestrians, and if this feature meet this need.

Target Acquisition

One participant expected a GPS to inform the user of when the target had been reached. As with other GPS this prototype announced the ending of the route, not with a verbal cue, but with the ending of beacon sounds and a short melody. The participants all seemed to understand the audio icon when they first heard it: *"Oh yes! Are we there?"*

5.3.3 Interaction

Different methods of interaction with the prototype was observed and talked about.

Motion

Two of the participants walked with a confident pace with the prototype, the others appeared more careful and trying in their speed. When they were uncertain about directions they all had a tendency to stop: *"Now I have to stop a little to figure out where it is."*. When they had found the direction of the beacon they again started walking. After some training, some soon manage to do turns and other route changes while walking.

One of the participants tried to search for the direction of the beacon by head turning even though the participant kept the body still. Since it is only the attitude of the iPhone that is registered, the participant was instructed to hold the head still and rather turn the body around. The participant soon grasped this restriction and did not make the same error later in the test.

Scanning

The spatial location of the audio beacon was sometimes reported to help the participants find the direction at once. When they were uncertain all participants seemed to adopt a method of scanning the surroundings by turning their body back and forth, looking for the snapping sound. This scanning behaviour appeared to end when the snapping sound was heard, and was used less with more training. When a participant was asked if the direction of the audio beacon was noticeable when the snapping sound disappeared, the participant said: *"Yes! No... I had to search. I tried to listen as well, and I tried a little back and forth, or from side to side."*

Another participant explain the search behaviour as a way of getting direction confirmation: *"Now it comes from the left, and then the snapping sound goes away. That is why I search for the snapping sound - to know that I am on the right track."*

One participant also adopted a scanning method where the body and head were held level, while the iPhone was turned around scanning for the snapping sound. The participant reported feeling this was a natural way of using it. This is the same technique used with the haptic pointer in **?** study.

On and Off

The possibility of turning the beacon on and off during a route was not implemented, but was mentioned by all the participants as a necessary feature: *"Yes! You had to. You had to absolutely."* One of the reasons mentioned, was to avoid being disturbed when listening to natural sounds from the environment. Another was the lesser need for directions when walking in familiar places: *"If I get into an area where I am familiar it is obvious that I would have turned it of. Then I would have turned it on again i such situations when I wanted to cross an open space"*.

5.3.4 External Influence

The test setting was affected by external influences that could have affected some of the results.

Headphones

While most of the participants did not speak of any problems with the earphones and was probably not affected by their presence, one participant struggled to use them. They did not fit in her ears and had a tendency to fall out. *"Yes, I hear if I hold them both. (..) Is it only in one ear I am intended to hear it?"* This clearly made it hard for the participant to use the prototype and experience the notion of spatial audio.

Noisy Environments

The test was conducted in a quiet neighbourhood away from traffic, but possibly disturbing noise was sometimes present. People that walked by and chatted, sometimes seemed to distract the focus of the participants. A street cleaning truck ones drove past during a test session. The noisy truck was only meters away, while the participant walked on confident. When asked, the participant reported to not manage to hear the spatial audio beacon in the noise, but the snapping sound was still audible. When asked how that could be, the participant answered: *"Yes, because it goes on another frequency"*.

Assistance

As earlier mentioned were the participants accommodated by the researcher at all times and assisted in finding directions when the prototype produced bad readings. This has to be taken into account when considering the confidence of some of the participants when walking with the prototype. If the prototype was to be used without assistance it was indicated that the participants would have been more concentrated and trusted the device less in choice of direction: *"But now I know that you are there, so.."*.

5.3.5 User Experiences

The concept behind the prototype was approached by the participants with positivism and they seemed to find it interesting and perhaps useful. The participants seemed to experienced the prototype differently, which produced different feedbacks.

Navigation

> *I should very much like to have one of those, if only in my neighbourhood. It had been much simpler to have something to walk after.*

When it comes to the navigation aspects of the prototype, the participants appeared generally positive. They seemed to accepted the prototypes limitations and errors as less important and could see themselves using a complete application on the same concept. When asked how the unstable directions of the prototype was experienced one participant said: *"I did not feel that this was so insecure. But, I should perhaps not get lost in the woods with it"*. Still other participants mentioned walking trips in the woods as the perfect environment for the applications where they would not need to be afraid of the application disturbing them in traversing traffic. The woods are also mentioned as a place as some want to be able to walk more in, where this application had the potential to help them.

> Researcher - *How did you experience the intersection back there where it gave you the wrong direction?*
> Participant - *No, it only says that you have to be a little awake yourself. But, if you walk at relatively known places, you can walk much faster when you have this one than when you do when you are uncertain and hesitant.*
> Researcher - *So it becomes an extra confirmation?*
> Participant - *Yes, you can feel much more safe and then you know that it has a little error, then you just got to try a little.*

The uncertainty due to technological limitations in the prototype caused problems when the participants tried to follow the directions given by the prototype. The participants did not seemed to become too irritated by this and seemed to learn to deal with it: *"You rather get used to not take it too literal then"*.

> Researcher - *How did you feel it went?*
> Participant - *I think it went very good! It was exceedingly.. No I think it was very nice. Funny!*
> Researcher - *Did you feel it gave you help where you needed it, or?*
> Participant - *Yes, if you look away from its margins of error.*

The participant who had troubles wearing the earphones complained about not being able to walk freely with the prototype: *"I can't walk in a smooth pace. I feel like I can't get to walk forwards, as when you walk normally."*. This could largely be attributed to a bad choice of earphones together with unstable orientation sensors, where the participant gave some positive responses later into the test: *"There the road is. This is not to believe, I do walk correctly."*.

Other participants did not have the same problems with the earphones and managed to use the prototype more freely. As the participant from the pilot test said after completing a route: *"It really worked, I got there without you saying anything"*. The same participant also experienced a moment where the potential of the concept really shined through all the technical problems. When walking in the middle of the street in a fast pace the participant reached an intersection and took a thirty degree turn and hit the next road with perfect alignment. All without slowing down, touching any road edges or receiving assistance from the researcher. The participant realized it, smiled and said: *It is very elegant. Now I handle turns!*.

Audio Icons

The audio beacon was often referred to as 'it', like: *"Where is it?"*. The participants reported that they felt it was positioned outside of themselves: *"I felt it was outside. I did. In the direction."*. Most participants found the sound sample used comfortable and nice, but the participant that had trouble with the earphones and felt the sound to be *"a little nerve-racking. Like you got.. I would have concentrated on walking."*.

Instead, all participants were positive to the snapping sound: *"I think the snapping sound was very funny, because it was like something new."*. One thing was the sound of it that the participants reported to be comfortable, another was the function it had: *"The snapping sound was very safe to have, because then I knew constantly that I was at the correct course."*. The snapping sound was also reported to originate from outside the head even though it was not spatial filtered at all.

Implementation

All the participants seemed to like the idea of having a GPS on their own cellphone, as one participant stated: *"It is a very nice way of orienting ourselves, by merely having a map-reader on the cellphone. That had been great."*. When asked about the use of such a GPS application in everyday life, one participant answered: *"Yes! I believe that actually. If I do not need to hold it while walking mark you, because that is maybe a bit too much trouble."*.

Due to technological challenges the prototype was designed to only work while being held flat. A more practical solution for use could be to have it hanging around the neck, something the participants agreed on. This will be discussed further in section 6.2.5.

Chapter 6

Discussion

In this chapter, we present our findings on the research questions stated in section 1.3. The findings are based on the multidisciplinary theoretical foundation from chapter 2, the knowledge gained through the development of the prototype as presented in chapter 3 and the findings from the interviews and field tests that are summarized in chapter 5.

6.1 Research question 1

> **Research Question 1:** What interaction design guidelines for the use of spatial audio displays in a mobile context can be inferred by theory

In this section, we will present the guidelines we have inferred during our research. The guidelines are mainly based on perception and design theory and their relation to each other with respect to spatial audio displays in a mobile context. Our experience from the development of the Sound Guide prototype and the following user-tests will be referenced in some of the guidelines.

6.1.1 Spatial Layout

Hearing is not the same as seeing. The medium itself and one's perception of it put natural limitations to how sound can be utilized in auditory displays. Interface design for visual displays is less limited with respect to layout. It is a natural frame the interface has to reside in, but inside this frame the designer is quite free to do as he/she pleases.

The Horizontal Arc

The limitations of the interface 'real estate' has to be accounted for in spatial audio displays. One vital aspect is the irregular distribution of *localization blur*. Localization blur can be described as the smallest change of spatial position of a sound event that the perception system manage to detect [**?**]. It can somewhat be compared to the aspect of screen resolution in visual screens. **?** base the recorded localization blurs on the smallest degree

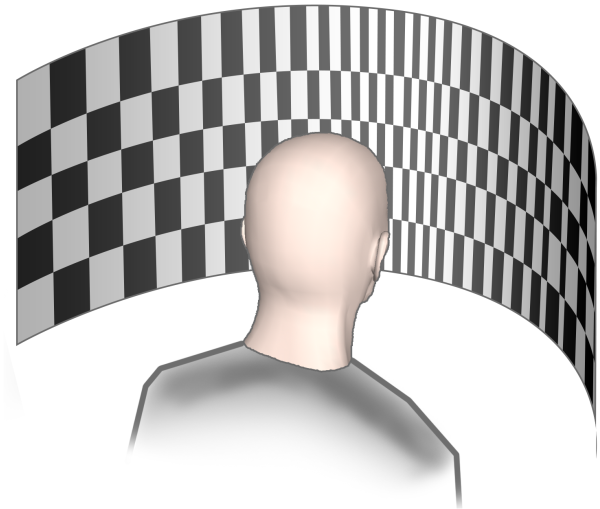

Figure 6.1: Visualized distribution of resolution of a virtual spatial audio display on a horizontal arc in front of the listener. Notice the uneven distribution of resolution in the horisontal and the vertical plane. This illustration is not to scale.

where at least 50 percent of the subjects managed to detect a change in localization.

The region that human hearing is most accurate is close to the center of the forward direction. Studies report localization blur in the horizontal plane to be down to about 1° in the center region, which should be quite impressive, with the complex methods of auditory localization taken into account. The localization blur increase from the optimal region in front of the user to the angles directly to each side, where the localization blur is larger with factors from three to ten times. In the rear of the subjects the localization blur again decrease and is larger than the front with a factor of about two [**?**]. Similar measurements of localization blurs in the median plane, show a minimum localization blur on elevation down to 4° for white noise and about $9 - 17^\circ$ for speech [**?**].

These numbers should be approached as the absolute limit of accuracy and should not be dependent on in actual applications. These recordings are all based on real sound sources, where virtual positioned sound sources at the moment reach a much higher localization blur that range from 10° to 20° [**?**].

The figure 6.1 gives an illustration of the distribution of localization blur with respect to the listeners head. Direction in the horizontal plane has the highest resolution with most sensitivity in the center angle. Resolution in the medial plane is less than in the horizontal plane and has also highest resolution in the center angle. The back side has a similar distribution, but is less sensitive. The dome above and below the listener follows to same horizontal distribution, but this sensitivity together with medial sensitivity blurs out when we reach the top and bottom side.

These findings lead us to the first guideline of this chapter.

> **Guideline 1:** Spatial audio displays should present elements, where the spatial location of the elements is vital, on a horizontal arc in front of the listener.

This guideline is not absolute. It should only serve as a recommendation if the concept of the spatial display can accept this limitation. Aesthetically elements need not be bound by this, but interface elements where the spatial location convey information should be positioned on the horizontal arc. It is possible to use elevation if the spatial filtering used support such cues, but the elevation should be restricted to only a few different levels and not be given too much importance.

The distribution of localization blur described above is the main argument for this guideline, where sound events that are positioned outside this front arc can be hard to localize for users and lead to confusion. This can be seen as natural constraints of human sound localization and should be seen to in interface design. This recommendation was first presented by **?** that referred to it as the 'Stage-Arc', and several other studies follow these recommendations [**???**, see].

Even though, horizontal localization is quite good behind the listener, the positioning of rear elements often lead to front/back confusion that is hard to avoid. Sounds in the back may also mask sounds coming from the corresponding angle in front. Some of these limitations can be overcome by head tracking, individual HRTF profiles and high quality headsets, but confusion and uncertainty may still occur. It will therefore be beneficial to rely on either the front arc or the back arc, and not on both. **?** design principle *"Design for error"* is much appropriate given the concern of front/back confusion and may be met by relying on a single arc in sound layout.

Relative Distance

With respect to the horizontal arc, *depth* could also be an important axis in interface design. Even though we are used to deal with distance to objects, the absolute distance to sound events is hard to determine from sound alone.

> **Guideline 2:** If distance of a sound source is to be used informally in a spatial audio display, it should be used relative to the distance of other similar elements present in the interface and not as an absolute informal state.

This is again only a recommendation and does not affect aesthetically use. Nor, distance used to accommodate spatial placement in a virtual environment where several aspects intervene to emulate spatial distributed elements. The important limitation in using distance is to avoid the expectation that users should be able to detect the state or distance to an object only by the intensity of the sound. Distance is partly determined by differences in intensity for sounds where the intensity is known [**?**]. We can easier judge the distance of speech over the distance of electronic tones because we have experience with the *normal source intensity* of the first.

Absolute intensity will anyway be hard to judge when using headphones, where the output volume can be different between users, and especially in a mobile context where outside noise can make sound intensity hard to judge. Highpass filters and room reflections may also aid in distance localization, but they do not eliminate the challenges with distance perception. Distance should be easier to judge when the intensity of one sound event can be compared to the intensity of other sound events that are present in the interface. When two distances are compared a larger difference is easier and faster to detect than small differences [**?**]. Drastic changes in the distance of objects should be audible where the memory of the earlier state is used as a reference point, but this impression of distance may diminish over time.

6.1.2 Sound Design

The choice of sounds used and the way they are presented have much influence on the perception of spatial audio displays.

Type of Sound

Several studies state that the use of broadband sounds, sounds that consist of a broad range of frequencies, are beneficial for sound localization [**?**]. A broad range of frequencies in the sounds presented makes it more easy for the brain to detect spectral cues, and several types of cues can then be used to determine sound origin more precise. One example here are the limitations of the duplex theory. If a sound source emit a pure tone below 1000 Hz, localization will be possible through interaural differences in phases, while intensity differences will not be detectable. If a pure tone of above 1600 Hz is presented, intensity differences will be audible, while phase differences will not be [1].

Sound localisation is possible for most audible pure tones, but sounds that consists of a broader range of frequencies are more beneficial. How much does it as to say? There seems to be little difference on the absolute minimum localization blurs with respect to type of sound stimuli in the horizontal plane. Clicks, pure tones, speech, narrow and broadband noise all seem to reach minimum localization blurs below 3° as can be seen in table 6.1. However, localization blur in this respect only deals with audible changes in horizontal localization of sound sources in ideal conditions. Elevation cues and front/back discrimination are more reliant on sounds with a broad spectrum [**?**] and virtual sound source positions may also benefit from it. [**?**] reported less reliable sound source localization when sounds with low frequencies were used and recommended using high

[1] Even if the human auditory system can not detect phase difference above 1500 Hz, it may still detect phase differences in the envelopes of the sound [**??**]. Sound icons with distinct start or stop envelopes may therefore affect localization performance positively. This effect is still much discussed and should be considered minor.

Type of signal	Localization blurs
Impulses (clicks)	$0.75° - 2°$
Sinusoids (pure tones)	$1.0° - 4.4°$
Noise (narrow or broad)	$1.4° - 3.2°$
Speech	$0.9° - 1.5°$

Table 6.1: A collection of results gathered from several studies as presented in **?**, p 39. As these numbers result from different studies that use slightly different methods, the localization blurs is therefore not fully comparable.

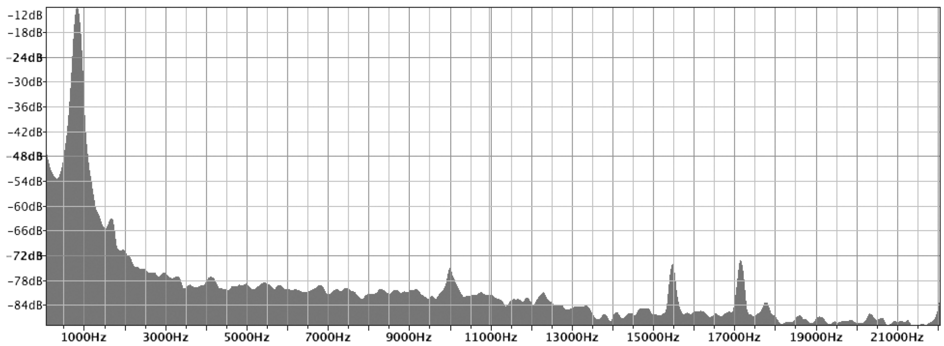

Figure 6.2: Spectrum analyses of the *Purr.aiff* sound clip used in the Sound Guide prototype. The graph shows that the sound used consisted of a broad range of frequencies.

pitched tones above 2000 Hz. Even though they can be used in some cases, pure tones should be avoided when sound source localization is vital.

> **Guideline 3:** Broadband sounds, sounds that consist of a broad range of frequencies, are beneficial for sound localization and should be considered.

Although broadband noise is most easy to localize based on the broad range of frequencies, there are many other aspects of sound design that also should be considered. When the beacon sound icon was chosen for the Sound Guide prototype, the spectrum and localization ability was evaluated. This was done by studying its spectrum analysis (see figure 6.2) and testing how accurate it was to localize. Affordance and 'pleasantness' was also given much priority in this evaluation. Does it feel natural to follow this sound, and is it pleasant to listen to over long periods of time? This aspects concern with mapping in interaction design where metaphors and analogies can be used to convey the function of interface elements.

Psychological cues should also be considered when choosing sounds and especially when using analogies and metaphors in the sound design. This is discussed further in section 6.1.5.

Presence and Duration

Users should be given time to get accustomed to the sound field. Externalization of sound sources becomes more prominent when users expect the sounds present and their position in space [**?**]. This can be achieved through a consistent use of sound events, and through the use of sounds with enough duration to permit satisfactory localization. It could also be beneficial to let sound sources be 'announce' in advanced to let users be aware of their presence before localization is needed [**?**].

> **Guideline 4:** Time should be given to let users become accustomed to the sound field, while static positioned sound sources with long duration should be avoided since they decrease the perception of sound source position.

Although sound localization may benefit from sound events with some duration, static positions of audio for prolonged durations should be avoided. If the auditory system is stimulated for a long period of time with the same signal, it will adapt and grow fatigue and the experience of localization will be reduced [**?**]. This effect may occur after a few seconds and slowly grow stronger until it reaches a maximum after about 3-5 minutes. **?** reported these problems in their prototype when their sound sources used seven minutes to move from one side of their 'Stage-Arc' to the other. Spatial audio in audible interfaces should be constrained to dynamic events, where the user itself affect the position of sound events, or for sound events with a relative short duration.

Order of Events

In sound environments with multiple sources of different spatial location, the first wavefront that reach the listener seems to have precedence in the interpretation of spatial localization. If the signals are similar and the delay is short, the first wavefront appear to mask or inhibit the later signals and only the first spatial location is perceived. This is called the *precedence effect* in psychoacoustics and concerns the order of signals in short timespan and their effect on perception [**?**]. This concerns mostly sound reflections that originate from other directions than the primary sound, but does also affect other sounds.

Since the spatial attributes of the earliest signal surpass the same attributes of late signals, it may be favourable to design spatial audio displays to give the most important element of multiple spatial elements the 'lead start' to aid in localization [see **?**]. It may also be wise to avoid presenting similar sounding spatial distinct sound events, as clicks or sound icons, without separating delays of about 100 ms or more [see **??**].

> **Guideline 5:** If several sound sources should be presented simultaneously, the most important element should be given a lead start, or the presentation of individual sound sources should be separated by 100 ms or more.

It should be clarified that this only applies when several sound sources is presented at the same time, where the duration is short and the sounds is of similar type.

6.1.3 Interaction

? claim that if the use of head-tracking is enabled in a realistic way, it will reduce the occurrence of front-back confusion and inside-the-head effects. It would also make it less necessary to provide individual HRTFs in applications. The ability to change the listeners space dynamically based on head-rotation, give the listener *"an improved sense of involvement, as they now perceive themselves as moving in an otherwise fixed scenario"*[**?**, p 383].

This augmentation of audio in an interactive environment can be seen as a good starting point for selling a concept with the right mental model. As human beings we are used to interact in a stationary environment, interface concepts based on the same idea should therefore be easy to adopt. It can be seen as relying on a standard of orientation human beings have grown accustomed to, during a lifetime of natural interaction.

Head tracking is beneficial in producing this effect, but body oriented tracking is sometimes sufficient. Especially in a mobile context, where the user is standing upright or walking, rotation of the body is effortless and natural. The test users in the Sound Guide case quickly learned to keep the head still and rotate their body to detect sound orientation. Instead of developing methods for detecting head rotation, a simple strap for hanging the device on the user's chest could be sufficient. The need for head tracking should still be evaluated for each application.

> **Guideline 6:** A spatial audio interface that lets users dynamically change the orientation of the virtual soundscape through head- or body-tracking is highly beneficial.

This enables users to interact with sound sources as they were real elements in the environment. As mentioned, this will reduce front/back confusion, give more realistic spatial perception and provide a solid conceptual model in the design.

6.1.4 Sounds in the Environment

Especially in the context of navigation applications for visually impaired pedestrians, the attendants to natural sounds in the environment are crucial for safe navigation. Headphones have been discussed for their negative effects in blocking ambient sounds, but the playback of audio can also block ambient sounds by producing masking effects [**?**]. Sounds used in auditory interfaces may hide other natural sounds that occur at the same time. This was mentioned in the user tests of Sound Guide prototype where users wanted to be able to turn off the interface when they encountered areas with street traffic, because they were worried that the sounds of the application would disturb their audible perception of the environment.

In contrast to this problem, there are also times when the sounds of the environment can mask out the sounds of the interface. This happened during a field test of the Sound Guide prototype, where a street cleaning truck came close. The user did not complain about the prototype disturbing the sounds from the environment, it was quite loud enough, but reported to be unable to hear the beacon sound over all the noise from the truck.

This can be seen as a breakdown in design, following **?** definition of the concept. Suddenly the information delivered by the interface is not present, and uncertainty may arise.

In a mobile context, these forms of breakdowns should be anticipated. Many methods can be used to combat this, like turning the volume up or using headphones with stronger insulation, but sound design may also be of aid. The snapping sound used to confirm front direction, proved to be clearly audible even in noisy environments. This may rise from its distinct character and broad spectrum.

> **Guideline 7:** When designing spatial audio displays the impact between natural sounds of the environment and virtual sounds from the interface should be addressed.

This is mostly a contextual concern where the priority of the natural sounds has to be compared to the priority of the interface sounds.

6.1.5 Realism or Enhancements

There is a large amount of interplay between the different senses in the task of perception. In the task of spatial localization through sound, people are used to be able to confirm the source localization through their vision [**?**]. If a subject are presented with a spatial sound filtered to a frontal location, the sound may be perceived to come from the rear since no possible source object is located in the visible field [**?**]. This types of 'reversals' in spatial audio listening is quite common, but probably, not as widespread in the case of visual impaired subjects. Similar reversal may occur if the sound source can be felt with the hands [**?**].

There is indications that sound localization partly is a top-down cognitive process that is largely hypothesis driven [**?**]. The central nervous system state hypothesis about the possible location of sound events and use the different senses to confirm or reject them. If the hypothesis is confirmed by the visual sense or the binaural characteristics of the auditory event, the percept of sound localization is formed and we experience the sound to have a spatial origin. Earlier experience may also affect this process [**?**]. If the recording of a helicopter or air plane is presented over headphones, the large majority of listeners will turn their head upwards in trying to locate the auditory event, even if the audio is binaural filtered to appear from below. Experience and expectations may affect the perception of location and should be considered when designing spatial audio displays. This can be attributed to the design principle *"Get the mappings right"* by **?**. The use of familiar, everyday sounds as analogies may benefit from spatial placement that coincide with each sounds natural placements in real life: Airport announcements comes from above, running water comes from below, thunder are far away, speech comes from eye level, bird twitter comes from high above, and so on.

> *Nonetheless, it would not be unreasonable to propose that spatial experiences that challenge or contradict natural experiences (or that are suggested by the visual sense) might lead to discomfort or dissatisfaction with the product.* [**?**, p 8]

What we psychologically perceive as natural put strong influence on how users react to spatial audio displays. These cues can also be used as an advantage. The developers of the *Papa Sangre* game decided not to use spectral filtering to simulate sound source elevation because they found psychological cues to be a sufficient clue [Paul Bennon, Chief Creative Officer in Somethin'Else, personal communication in November 2011]. Instead of making their engine more complex and processor heavy they used sound design (helicopter sound comes from above) or dialogue ('Look up!') to give the perception of elevation of sound sources.

> *..authentic reproduction is rarely required. More often the technological task is to enhance or augment acoustical scenarios rather than to replicate them authentically.* [**?**, p 374]

To reduce front/back confusion in the Sound Guide case, a lowpass filter was put on the sounds originating from the rear. This occurs naturally by the shape of the ears and was probably applied in the HRTF filters used on the sound icon as well. Preliminary testing still revealed that it was hard to detect if sounds originated from the front or the rear with headphones and sound filtering used. Based on this we decided to *enhance* the traditional psychoacoustical effects by applying a stronger lowpass filter that would be audible even on cheap headphones. This filter received critique to be too soft to indicate anything, and it can be argued if this is a good way of resolving front/back confusion, but it illustrates a method of design where realism is not necessary the goal.

> **Guideline 8:** Natural psychoacoustical cues in sound localization should not be seen as rigid limitations on design, but rather as a toolbox in achieving more usable and intuitive designs.

Future research on the use of spatial audio in auditory display design will hopefully give a clearer picture on how these effects can be used most efficiently.

6.1.6 Individuality in Perception

There is a large diversity of ear shapes and head and torso sizes. To give optimal spatial audio emulation for each application user, individual HRTF profiles would be most beneficial. It could be likened to get your vision checked where each user would have to go to a specialist that could construct a personal HRTF-profile based on recordings from each ear. This is not that easy to achieve commercially, especially when the use of spatial audio is not that widespread.

A more simple solution would be to provide a small range of profiles the each user can select from. This will not provide listening experience that are optimal in all cases, but may provide HRTF-profiles that are a close enough match for most user for practical use. The research project *The Sound of Football* provided ten different profiles their participants could choose from. The *Papa Sangre* game provides five different profiles, categorized through the sex and the hair length of the user.

There have been much research in constructing general HRTFs that provide a good spatial perception for most users. This has proven not to be that easy to achieve. Even though, a general HRTF profile seldom produces optimal sound localization, it may in many cases give sufficient spatial perception to be used in applications. A general HRTF profile was used in the Sound Guide prototype, and seemed to provide sufficient sound localization in the tests given.

Inadequate HRTF profiles produce more front/back confusion and it degrades sound source localization. Especially elevation cues are hard to use through general HRTF profiles. This challenges can be combated by making more simple and robust interfaces. Head or body tracking may minimize the problems with front/back confusion and are recommended in such situations. Supplementary features like heading confirmation can also be beneficial in cases with general HRTF profiles.

> **Guideline 9:** If individual HRTF profiles is not provided, care should be given to minimize the implications of inadequate sound localization.

Even though general HRTF profiles can be sufficient in most cases, the application should be designed to accommodate the cases where it is not. As **?** encourage: *"Fixed solutions will invariably fail with some people; flexible solutions at least offer a chance for those with special needs"* [p 164].

Effects of Hearing loss

One of the participants in the Sound Guide case used a hearing aid in her right ear that was removed before testing. The participant was used to walk without it, but stated to have mild hearing loss on the right side. In the controlled testing environment, this was not apparent where the participant score of accuracy was in some cases better than for other users without hearing loss. Testing in the field, on the other hand, revealed that the participant in many occasions experienced the sound beacon to disappear when it changed position to the right hand side of the participant. This caused stress where the users sometimes needed to turn around completely before the position of the 'disappearing' beacon was found. This observation indicate that the use of spatial sound in audio interfaces may produce problems for users with imbalanced hearing loss. **?** confirm that the accuracy of sound localization decreased on people with asymmetrical hearing loss. Symmetrical hearing loss had little effect on sound localization accuracy if the loss where below 30-40 dB. Age related hearing loss may therefore not affect the use of spatial displays noticeably.

6.1.7 Implementation

The hardware and software support on the device the spatial audio display is developed for put certain limitations on the design of the interface. If the spatial audio is filtered in real time at the device, the number of simultaneous sound sources is often limited. The range of acoustic cues that are supported is also up to change. Real time spatial audio filtering is normally processor heavy and care should be given to minimize the footprint of the application. This is especially important if the filtering is done on the software level.

If real time filtering of spatial audio is not supported, a prefiltered approach can be chosen. This was done in the Sound Guide case, where complete sets of spatial filtered sound clips from different locations was made during development. These clips were then played back by the application to represent different locations. This approach put strong limitations on the sound design, where static sound icons need to be used and their possible positions must be foreseen. Dynamic content by synthetic speech is not possible to position spatially following this method. The number of simultaneous sound sources is also limited.

It is possible to externalize the spatial filtering of dynamic content through the network. While the mobile device serves as the terminal for the user, a server or personal computer may produce the spatial filtered audio and transfer this audio to the device by the wireless network. Even if this approach can is useful in some situations, it will produce a high dependency on a stable and fast network connection, and it would probably produce noticeable delays in interactive applications.

Headphones for Spatial Audio

A study on how visually impaired users evaluate different output features for a navigational aid, the use of headphones was rated as one of the least accepted [**?**]. Those with some vision left showed more acceptance to the use of headphones, than did the blind participants. This may indicate that the blocking of ambient sounds used in orientation is the main concern for avoiding the use of headphones.

The best headphones to use for spatial audio should have a flat frequency response that do not alter the subtle cues used in sound localization [**?**]. Headphones labelled with 'bass-boost' or 'noise reduction' may not be ideal for this type of use, since they are designed to alter the sound that is reproduced. The original Apple iPhone™ earphones (version MB770G/B) used in the field tests is not ideal on this subject. The frequency response curve has a dip in the high and the low frequency range, and the distortion is quite bad. This may have affected binaural cues that rely on the very low and the very high ends of the spectrum.

One of the advantages on using the Apple™ headphones in respect to outside navigation is the low level of isolation. As stated in one of its reviews: *"Their isolation capabilities are, quite simply, nonexistent"* [2]. This prove to be important for visual impaired users when navigating outside, since they have to rely on their perception of ambient sounds when navigating themselves in the environment. Headphones that block natural sound from the outside could lead to dangerous situations for people relying on that sense for navigation.

An alternative set of headphones with the same low level of isolation would be the popular Koss Porta Pro™ headphones that have a better frequency response distribution and less distortion. That set of headphones may have been a better choice for the field tests when seen in retrospect.

Another approach would have been to use bone conductance headphones that do not interfere with the outside ear. Such headphones may not give the best sound reproduction, but **?** have showed that such headphones still produce decent results for auditory display interaction.

Headphones that encapsulate the whole ear often have a high quality reproduction of sound. The *Beyerdynamics DT 770 PRO 80*Ω headphones used in preliminary tests have an even frequency response curve and barely any distortion. However, **?** points out that such headphones may influence sound frequencies above 1 kHz due to resonance inside the headphones and will distort psychoacoustical cues that reside in the higher spectres. How much this has to say compared to other headphone issues is not mentioned, but it could indicate that high quality in-ear headphones is the best choice for spatial audio reproduction. Such headphones avoid influence of the outer ear on the spectral cues, and may have a more direct control over the sound that reach the inner-ear.

> **Guideline 10:** High quality headphones with a flat frequency response are best in reproducing spatial audio, but acceptable reproduction can also be achieved with poor quality headphones in most cases.

High quality headphones are perhaps the best choice when it comes to enabling subtle binaural cues in realistic reproduction of spatial sounds. Still, it may not be necessary in every use-case of spatial audio display. As with the example of blind pedestrians above, there may be other needs or practicalities that make it not feasible to expect all users to wear expensive headsets. Especially, for applications distributed on smartphone platforms where the developers have little or no influence on the type of headphones the user choose to buy. The diversity of headphones in the consumer market is wast and applications that aim for that market should be usable even with cheap headphones. This can be achieved with testing.

---

[2] `http://www.headphoneinfo.com/content/Apple-iPhone-3G-S-Headphones-Review-899/Isolation.htm` (Accessed 16. July 2012)

6.1.8 Summary

In this section, ten interaction design guidelines have been presented. The use of spatial audio in auditory displays in a mobile context is restricted by psychoacoustical aspects of the human perception. These restrictions influence the design of spatial audio displays.

The position of virtual sound sources with respect to the listener has certain limitations. The optimal area to position virtual sound elements is on a horizontal arc in front of the user. Positioning elsewhere is possible, but can lead to uncertainty and bad localization abilities.

The type and presentation of sounds also affects sound localization. A broad spectrum of frequencies in the sounds used is beneficial in producing robust virtual sound positions. The listener should be able to have time to get accustomed to the sound field, but prolonged sound playback with spatial placement should be avoided.

User interaction influence how sounds are perceived. The use of head- or body-tracking is highly beneficial, since it reduce localization errors and provide a natural way of interaction with the virtual display in a mobile context.

Authentic reproduction is not always the main goal, where psychoacoustical cues can sometimes be used selectively to achieve certain effects.

Best sound source localization is achieved using individual HRTF-profiles and high quality headphones, but general HRTF-profiles and poor quality headphones are in some cases sufficient if the interface is designed with these restrictions in mind.

6.2 Research question 2

> **Research Question 2:** How was the prototype that utilized spatial audio on a smartphone for directional guidance received by visually impaired test users and what can be done to improve the concept in later development

To answer this research question the results from the user tests will be compared to the guidelines inferred in the previous section, the related theory and the results from related studies. The informal field tests conducted makes it hard to make conclusive findings, but possible causes and solutions will be discussed on the aspects of the prototype that was found to produce challenges for the users.

6.2.1 The Sound Guide prototype

The Sound Guide prototype was developed as an application for a commercial smartphone platform that utilize spatial audio for directional guidance in route navigation for visually impaired pedestrians. The resulting prototype focused on route guidance where directions were presented through the location of spatial audio icons. Compared with most electronic O&M aids for visually impaired users this is a quite limited set of features. Creation of personal landmarks with additional information, editing capabilities of routes, 'look around' functions, automatic route creation, address searches, route sharing and route merging would all have been relevant features to develop, but this went beyond the scope of this study. This section will discuss the design of the route guidance features implement based on the results from the user-tests and the relevant theory. Challenges will be pointed out, and suggestions for improvements will be presented and discussed.

It is vital to point out that the prototype was designed as an orientation aid, and not a mobility aid. The user needs the application was hoped to meet was providing updated heading information that the user could use to update their direction with respect to recorded routes. This was intended to 'keep them on the route' but not necessary 'keep them on the road'. The second, proved not to be achievable given the limited accuracy of the sensors used, where larger scale orientation was seen as a more suitable goal than small scale. Orientation aids will still be able to facilitate mobility to a certain extent in the sense that users can get familiar with the environment and know what type of terrain to expect. Still it is orientation with respect to route guidance that will be the topic of interest in this section.

6.2.2 Directions by Compass

A fair amount of route guidance aids has been developed for visually impaired users. However, of our awareness, no such commercial products utilize compass sensors to provide directional information spatially. **?**

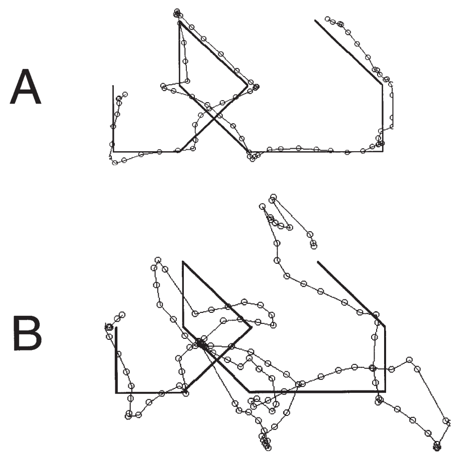

Figure 6.3: Examples of two different tracks of route performance. One with an angular bearing with compass sensor (A) and angular bearing based on GPS fixes (B). Illustration from **?**, p 200.

tested different display modes for directional guidance on completion time and travel distances of abstract routes. Virtual mode (similar to the spatial audio beacon used in the Sound Guide) produced the fastest route completion time, and produced most accurate route traversals. Other display types tested was left/right/straight indication, angular bearing ("left 80") and a no compass option that gave angular bearing based on changes in GPS-position. As illustrated in Figure 6.3 guidance by compass sensors was in some cases superior, and produced mean travel distance and completion times that were twice as much as the virtual guidance mode [**?**]. Subjective ratings also showed that five out of eight participants rated virtual mode best, where the no compass mode got less favourable ratings.

These results were obtained from a test field that were free of obstacles where the audio cues was the only reference of route orientation. Route navigation in more bounded areas is bound to be more suitable for traditional GPS guidance since more orientation cues are present in the environment to guide the pedestrian. Here, the visually impaired users can follow the edge of the side walk through a white cane and be able to travel routes based on cues in the environment with more accuracy.

Spatial direction information can still be beneficial in these environments. **?** found that visually impaired participants accessed directional information more when they navigated semi-bound areas like a park with curved paths, then in bound areas like streets with straight side walks. This can be interpret to show that directional information from aids is accessed when the environment does not provide enough directional cues for orientation.

Route 'start-up time' may also benefit from compass sensors. One participant from the Sound Guide case had experience with using a Trekker Breeze GPS and was used to get directions based on GPS-tracking: *"When you walk in a direction you will gradually figure out if you walk south, east, north* *or west"*. This was especially apparent at the start of routes. The participant was interested in a device that could give this directional information without having to guess the heading to walk in and wait for the verdict. This would save the participant time and simplify the task of orientation.

All participants from the Sound Guide case field tests were positive to get updated direction information and stated interested in obtaining the same functionality in commercial systems.

6.2.3 Sound Design

Most auditory displays use synthesized speech as the primary way of communicating information. The Sound Guide prototype instead used abstract audio icons for all interface elements and the order, choice and spatial positioning of these elements to communicate information. This has several benefits: Non spoken sounds can easily be ignored and distract less from navigation [**?**]. Detection of spatial perception of sound sources require less cognitive and memory loads, compared to verbal direction cues [**?**], and is also faster to detect [**?**].

A critique of non verbal interfaces is the ambiguity of the abstract icons used. Users have to learn and understand the meaning and relations of audio icons before they can utilize such interfaces effectively. This put a stronger responsibility on interface design. Few standards in sound design have been developed to accommodate these problems. Until such standards emerge, each application has to provide a consistent conceptual model for its own use that provide good mappings and easy learning.

User feedback from the Sound Guide case shows that people have different preferences when it comes to sound design. It may be beneficial to provide customization abilities in commercial auditory displays [**?**].

Heading Confirmation

Pilot testing of the Sound Guide prototype revealed that users felt uncertainty in determining when the spatial audio beacon was straight ahead. A heading confirmation was added. When the heading of the user was less that 20 degrees of the route direction a snapping sound was presented. This gave an audible indication that the user was walking in the right direction, which was not dependent on the seemingly fragile perception of sound localization.

All participants seemed to like this heading confirmation and was observed to depend upon it as the main indication of directional information. All participants that tried the prototype with the heading confirmation, adopted a technique of *scanning*, where they often in times of route changes rotated their body back and forth in search of the snapping sound to confirm the new heading. This can be compared with the techniques used with the haptic pointer interface in the UCSB Personal Guidance System by **?** where users used a pointer to find the direction that produced audio feedback.

Since the snapping sound used a slow repeat rate of once every other second, the sound element proved to be easy to miss by rotation. This is probably why the users often went back and forth many times when scanning for the sound. Later versions of the prototype, played a snap every time the heading zone was first encountered during rotation, to prevent this behaviour. This fix was never tested on users.

What about Sound Localization

The main concept of the Sound Guide case was to utilize spatial audio filtering to convey directional information. The introduction of the snapping sound (not spatially filtered in the tested versions) did in many ways remove the need to localize the spatial placement of sound sources. Users could use scanning techniques to find the heading confirmation and follow routes, without thinking about spatial location at all. This tendency was not anticipated by us and made the results harder to analyse with respect to spatial audio. The test could be said to show limited internal validity on the use of spatial audio, since alternate directional information was presented. Did the participants perceive the heading of the sound beacon when snapping sound disappeared? Did they use this information in finding the next route heading? Participants reported to be able to use the perceived sound localization in detecting new headings, but observations showed a strong tendency of scanning behaviour that may indicate that this was not used in a great extent.

The pilot test conducted without heading confirmation showed that spatial audio can be used to communicate directional information. The test of sound localization that was done indoors also showed that the participants were able to localize virtual sound positions. Still, the quiet indoor setting of this test reduce the external validity when it comes to the appliances in pedestrian navigation in outdoor environments.

[?] used a virtual beacon in their UCSB Personal Guidance System prototype that did not give any heading confirmation beyond the spatial position of the sound source. This display mode proved to be faster, more accurate, and overall received better user ratings than the haptic point interface mode. No uncertainty of heading was reported in this study. So is it really necessary to use heading confirmation in spatial audio route guidance?

This may be ascribed to the quality of the audio filtering in providing stable perceptions of sound localization. The spatial audio used in the Sound Guide prototype can be improved on several levels. Individual HRTF profiles or real time filtering that does not fade between states, could produce better sound localization that makes heading confirmation redundant. Better quality headphones and more authentic sound reproduction may also be of aid.

Even though heading confirmation is not strictly essential, it has several benefits. It can be seen as a way of applying *redundancy* where it provides several ways of accessing the same information: either through spatial location, or through phonetic qualities. It can be seen as a design aid in that it subtle indicates that it is something special about one heading with respect to the others. In this way, it makes the 'right' choice of action more prominent. It also makes the interface *degrade gracefully* [3] when noisy environments, hearing loss, or missing headsets make spatial location imperceptible. This makes the interface more robust and perhaps increase the user experience. Users do not need to be uncertain about walking in the right direction or not.

This happened with a user with the UCSB Personal Guidance System prototype. After he had been walking past a waypoint, he did not notice it until the verbal distance measure started to rise instead of fall. An audible heading confirmation could perhaps make this error less likely to happen.

No Sound as Error

The heading confirmation was designed to be present when the user walked in the correct course. If the user deviated from the course, the audible confirmation disappeared. This can be characterized as a *no sound on error* design if we think of the snapping sound as the main audio component. The disappearances of the snapping sound often made participants stop. Even though the still present spatial audio beacon should provide them with the direction, they seemed to look for the 'missing' snapping sound. *"Where is it?"*, were sometimes said in such situations. Could it be that habituation made them not aware of the audio beacon?

In the state of course deviation, we want the user to use the spatial location of the audio beacon to orient themselves. To give more focus to the use of this sound, it could have been interesting to change the sound design to *sound on error* instead. When the user walk correctly the interface does not interfere, when the user deviates the interface becomes audible. This would provide a prominent recommended action in the case of breakdowns. It would also reduce masking effects and perhaps make a less intrusive electronic aid. **?** has also reached this conclusion with respect to GPS-guidance. The question concerns when the attention of the users should be focused on the interface and when it should be focused on the environment.

The flip side is the loss of feedback. How does the user know that he/she walks correctly or if the device has run out of battery? Distance will also not be communicated through the beacon and provide less spatial updating to help in orientation. Future testing of this concept is needed to explore the effects of such a design choice.

Front/Back Confusion

Little problematic front/back confusion was observed during the field test of the Sound Guide prototype. Much of this may be attribute to the reliance on heading confirmation. Preliminary testing without the snapping sound

[3] *Graceful degradation* is a term used within web development where functionality is designed to degrade to use older techniques if new techniques is not supported by the browser.

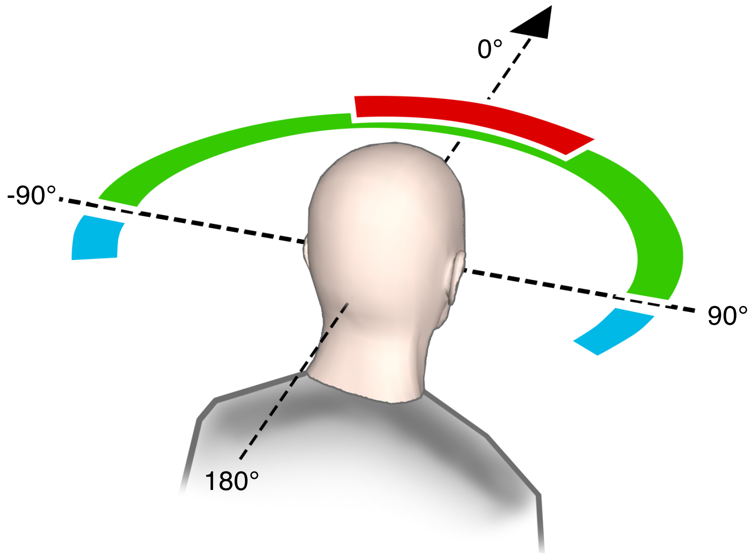

Figure 6.4: Visualized sound design where rear placement are represented as static *left* or *right* icons to comply with *Guideline 1* and avoid front/back confusion.

revealed much front/back confusion. The users usually resolved it through rotation, but uncertainty was introduced before this. Low pass filter was applied on the audio beacon in the rear hemisphere of the listener, but this did not resolve all front/back ambiguity. This design choice was also criticized by external supervisors on Media Lt to be too soft a cue to provide sufficient distinction between the front and rear hemisphere.

As argued for in section 6.1.1 the horizontal arc in the front hemisphere is the best area to present spatial information. Could front/back confusion be resolved by relying only on the front arc for spatial placement? One possible implementation following this design could be to simplify all rear placement to either left or right indication. This could be seen as providing artificial constraints to the interface make errors less likely to happen and the recommended action more clearly communicated. Especially in pedestrian route navigation, the goal when presenting waypoints in the rear of the user would be to get them to turn around. Accurate spatial placement might not be that relevant for these tasks, where accurate orientation is not that necessary. The sound design could then be seen as in figure 6.4 where rear placement is presented with a new sound icon, perhaps a low frequency sound with a diffuse rear left/right location or a just a verbal cue of *"Turn left"* or *"Turn right"*.

Distance

In route guidance aids, the distance to the next landmark or waypoint is a useful type of information. This helps users foresee their next move and help them orient themselves in the environment. The UCSB Personal Guidance System prototype by [?] presented distance information to the next waypoint as verbal cues every eight second. This rate coincided with how often the participants reported to want to hear such information.

The Sound Guide prototype developed in this study communicated distance through changes in sound intensity, as in natural environments. This was primarily evident when two waypoints was presented, and sound intensity could be compared between them. Distances above 50 meters produced a stable low volume (5% of full volume), while the intensity grew linearly while approaching the waypoint. This approach was also chosen by **?**. **?** used a similar approach where the rate of sound pulses indicated the distance to the target. Such audible representations of distance are used as reversing aids on some cars.

It can be argued that the intensity approach to distance presentation is too subtle and does not make distance predictions accurate enough. The user will be able to percept that the waypoint is getting closer based on the growing intensity, as reported in user tests, but will he/she be able to predict the distance to the waypoint? The verbal approach may have an advantage there. As **?** design principle *"Make things visible"* communicate, the users should be given feedback they can use to evaluate their actions. More concrete distance information could be one such feedback.

? found realistic distance perception through spatial audio to be a challenge. Their first prototypes used intensity difference to convey distance [**?**, see], their next prototype used both methods combined [**?**, see] and their last article did not mention intensity difference at all, but relied on verbal cues [**?**].

Turn Indication

When a waypoint was within one meter of the users position, or if the user had walked past it, the direction to the next waypoint was presented. Low accuracy of the device's GPS-position (5 meters radius of uncertainty reported by the device) made turn indication unstable and not reliable. Participants often experienced that the prototype instructed them to take turns several meters from the relevant intersection. This proved to be problematic because even though the participants got used to this form of error, they could not always foresee when the device turned too early and when it was correct. This produced a situation of breakdown in the design where the instructions from the application did not match the environment and no obvious method was present to resolve it.

Seen from the developers perspective a 5-meters radius of uncertainty on the users location, can be likened to an imaginary sphere with a diameter of 10 meters. The developer knows that the user should be located somewhere inside this sphere, but does not know where. So when the center of this sphere is located at an intersection, it is impossible to know if the user is right there, five meters past or has yet five meters to walk.

[**?**] used a high quality GPS-receiver with a large antenna to reach absolute accuracy shorter than 1 meter for their prototype. The small GPS-receiver in smartphones today can not match this accuracy, and probably have to wait for the development of new positioning systems [4] to increase their accuracy. External GPS-receivers [5] may also be connected to smartphones to provide better positioning and should be a recommended for these types of applications on this device.

Even though the best solution to inaccurate turn indication is higher positioning accuracy, interface design may also aid in making this less a problem. First task would be to make this uncertainty visible in the interface. Instead of making immediate turn changes, the interface should gradually communicate that a turn can be made when the user is two times the radius of uncertainty away from the intersection. This is where the front edge of the imaginary sphere mentioned above touch the intersection, while the user still can be positioned close to the rear edge. Lets call this the zone of uncertainty. Experienced visually impaired pedestrians are used to detect intersections based on aspects of the terrain (ending side walks) and the ambient sounds (traffic in front). One possible design could be not to present the second waypoint until the zone of uncertainty is reached. The introduction of the second waypoint will inform the user that a turn is coming and communicate which direction it is. At the same time, it presents the original direction. The two waypoints can be argued to communicate the ambiguity in the situation and let the user decide which waypoint that is best to follow at the moment based on cues in the environment. This form of ambiguity can be said to avoid breakdowns, because one of the directions presented is always right and the user therefore have distinct options to consider in resolving it.

This design combines turn indication with the earlier direction to come functionality. Testing is here needed to reveal if this solves the users uncertainty and communicate the ambiguity in instructions correctly.

Handling of Orientation Instability

The prototype suffered problems in determining position and orientation that affected how the route was perceived. Participants were often instructed to walk out of the road, or follow roads they should not follow. These problems can be attributed to the accuracy of the hardware sensors, but interaction design can still be used to approach some of these problems. A relevant question concerning these problems is: How can the application foresee its own instability? One goal with that kind of information, would be to let the applications uncertainty be visible to the users, so they could know when to trust the device and when not to do it.

Inaccurate position is discussed in the previous section with respect to this problem. Another vital problem is the unstable compass sensor that sometime showed completely wrong directions for longer periods of time. Several attempts were made to provide a stable orientation by letting the

[4] The positioning systems GALILEO and GLONASS may replace or be used in combination with GPS to give higher positioning accuracy in future mobile devices.

[5] An example here is *Bad Elf GPS* that increase the positioning accuracy of iPhones to 2,5 meters. See: http://bad-elf.com/products/gps/

different sensors compensate for each other's weaknesses, but this was not achieved with sufficient stability.

One method of making this general inaccuracy visible was to display heading confirmation with little accuracy. The snapping sound was displayed if the users orientation was measured to deviate less than 20° from the heading of the next waypoint. This gives a 40° zone of uncertainty in the display of direction that can be said to accommodate the inaccurate detection of heading. This put the user in charge of finding the right way inside this zone. Even if the road should deviate 20° from the heading provided, the user would still be given heading confirmation while following the road, and the inaccurate orientation would therefore not produce uncertainty. This zone of uncertainty could be changed according to the detected accuracy of the device. It may be beneficial not to provide directional cues when the accuracy is not precise enough to resolve choice of roads in traditional intersections.

If the sound of the beacon is perceived to originate from far away, it may also help in making the direction appear less accurate. This could be inferred to indicate that the application is not accurate enough to provide the avoidance of obstacles, where it is more suited to help in giving the general direction towards the far reaching destination of the user.

Ambient Masking

Masking effects happen when clearly audible sounds gets inaudible by other sounds in the environment. This has nothing to do about the blocking of sound by headphones, but rather the blocking of sounds by other sounds. In the case of the Sound Guide, the presented sound icons may make ambient sounds inaudible. Many visually impaired participants complained about this in **?** study, and participants in the Sound Guide case also stated concern about that effect of audible displays.

Several methods can be applied to combat this. The use of spatial audio is one, where the spatial attribute of the sound help the human perception system to separate it from other sounds in the environment. This can be attributed to the Cocktail Party effect described in section 2.2.2. Intensity and duration of the sounds used may also reduce the problem of ambient masking. If the sound design produces less sound, it interferes less. A last method will also reduce the occurrence of this problem. If the users easily can control when the interface can present sounds, and when to be silent, the users themselves can remove the chance of ambient masking in environments they want to attend to fully.

6.2.4 Audio Filtering

The spatial audio in the prototype was filtered using a general HRTF profile. Even though most participants reported to experience an out of the head sensation of the audio icon, the spatial position of audio seemed harder to detect in practical use while navigating outside. Individual HRTF profiles would probably have improved on the sound localization and externalization of sound events. Real time filtering of spatial audio on the device would also have been beneficial in producing natural reproduction of spatial position.

The sound was filtered to emulate free field environment where no room reflection or reverberation were presented in the sound field. This was based on theory that describe sound localization to be most accurate in this form (see 2.2.1). **?** used reverberation in the audio beacon in their USCB Personal Guidance System. This was done to make the sound to be perceived as more realistic. **?** state that reverberation aid in creating out-of-the-head experiences. Therefore, reverberation may be beneficial in producing a more realistic sound field. **?** used reverberation with low intensity to keep the reproduction realistic, while at the same time minimize the influence on sound localization. It may be that ground reflections also could add to the realism, and perhaps give stronger distance cues. More research is needed to determine the best filtering of sound for this use.

Is Spatial Audio Necessary

This work has shown that it is a large set of natural constraints with the use of spatial audio and a large set of considerations that have to be taken into account. It is natural to question if the advantages of using spatial audio in interface design outweigh the disadvantages.

A simplified spatial positioning can be achieved with using regular stereo panning. [**?**] indicated turns in their route guidance application by changing the balance of sound intensity from side to side. This proved sufficient in guiding bicycle messengers to their destinations. [**?**] and **?** did also provide directional information through stereo panning for use in GPS-navigation with decent results.

Preliminary tests conducted in this study (see section 5.1) have also shown that sound localization is possible through the use of simple stereo panning (down to 18° when presented multiple times). Most people state not to experience these sound events to originate from outside their head, but rather from some place between their ears [**?**]. The participants in our work describe the task of sound localization of stereo sounds as balancing the intensity between their ears to determine the direction. Perception of external placement is therefore lacking when stereo panning is used in stead of spatial filtering.

One advantage of spatial audio is the natural localization of sound source position. No techniques are necessary to learn because sound source localization is used in every day life. It relies on perceptual rather than symbolic processing which requires less training and instructions [**?**]. Front/back ambiguity is easier resolved by spatial audio than with stereo panning. Elevation and distance is also more precisely emulated through spectral cues. Stereo panning should be seen as a crude way of giving sound sources spatial positioning, where spatial filtering is the better choice.

There is also studies that show spatial audio to be the best choice when it comes to avoiding masking effects in a mobile context: *"Monophonic and stereo sounds as primary navigational senses should be avoided since using them may block environmental sounds."* [**?**]. This can be attributed to the cocktail party effect that among other things use the perception of sound source position to discriminate between sound sources.

The work conducted by **??** have also shown that virtual displays that use spatial audio outperform other auditory displays when it comes to route guidance. One advantage in using spatial audio in route guidance is that sound origin is communicated by the sound itself and does not require verbal descriptions or search behaviour **?**. This saves time and make the spatial interface more efficient and perhaps more natural.

It will be interesting to see what aspect of learning can influence the use of spatial audio. Even though novice users do not utilizing the full potential of sound localization and sound source discrimination in spatial audio displays, it is possible that experienced users will be able to operate much more complex displays than what have been shown in preliminary studies.

6.2.5 Implementation

The implementation of the prototype has combated the limitations of the proprietary smartphone platform *iOS* that run on *Apple iPhone* smartphones. The application was developed as a standalone application and has the ability to be published on the online market place *App Store* and reach millions of potential users. Other smartphone platforms could have been chosen, but iOS was chosen for this study based on its solid reputation in the visually impaired community and its technical abilities. We had positive experiences developing for this smartphone and with the other technologies used in the prototype. The sound library *FMOD* proved to be powerful and applicable. The spatial audio filter *Panorama 5* provided excellent control of sounds with respect to interface sound design, but ideally a real-time solution for the iOS platform would have been most beneficial. This would among other things allow for the use of spatialized synthesized voices and make it easier to implement individual HRTF profiles. We are waiting for the smartphone market to mature on the field of spatial audio, and will keep an eye out for the coming release of the 3D audio engine used in the *Papa Sangre* iPhone game (see 2.4.1).

Use of Smartphones

There are many advantages in the use of smartphones as orientation aids. **?** argument for the use of smartphones [6] based on their low cost (from both consumer's and developer's standpoint), technical abilities and mobility. Battery consumption is also mentioned as an important requirement for M&O aids were the device needs to be used for a longer period of time on battery power. This can be seen as a negative aspect of smartphones that

[6] In the article referred to as PDAs, as advanced mobile devices were called at that time.

serve many other functions apart from navigation and therefore will not last as long as more dedicated devices.

As mentioned earlier the positioning sensors used in smartphones are not necessarily the most accurate that exist on the market. This put an advantage on the potential of dedicated devices when it comes to precision. The compass sensor also proved to be less accurate in the device tested. These problems may be combated through combinations of other sensors and interface design that take these uncertainties into account.

The participants from the Sound Guide case all used cellphones in their daily life, but not all had tried more advanced smartphones like the iPhone. They were all positive to the idea of having an orientation aid at their cellphones. That would result in fewer devices they needed to carry around, and the use of traditional smartphones was stated not to stigmatize them as a group based on the devices they use. As **?** point out *"any M&O aid should not highlight the blind user as a person with disabilities"* [p 385]. **?** share also this concern by stating that the use of traditional phones would *"avoid the stigmatizing effect of products labelled "for disabled people""* [p 1230].

One of the visual impaired test users stated concerned of how the design of devices used in public, affected how the user was perceived by others. The user perceived some assistance aids to be so ugly that the user did not like to use them in public. The user was experienced in using the Trekker Breeze GPS, a handheld device, but rated it as bulky and impractical to carry based on its size. This device was sometimes noticed by others in the public as an unusual aid to use. The user would not mind carrying a cellphone around her neck as long as the device looked like cellphones other people used and therefore did not separate the user from others.

Headphones

The participants showed concern in using headphones while walking outside. They appeared not to have too much trouble with perceiving ambient sounds during the field-tests, but this problem can still not be ignored when developing orientation aids with binaural audio. **?** reported the same responses from their visually impaired participants. Headphones were reported to block the higher frequencies of the ambient sounds that were vital in orientation. To be able to use binaural audio without interfering with the participants perception of ambient sounds, they developed a novel headphone that delivered audio through thin air-tubes placed in the ear canal. Listening test with self reporting gave good scores for these headphones prototypes based on low blocking of ambient sounds.

Such specialized headphones may be the best choice, but we have not managed to find such headphones currently on the market. Mounted shoulder speakers have also been used for spatial audio displays, and should produce no blocking of ambient sounds apart from masking [**?**]. Headphones like the Apple headphones used at the field-test, and the popular Koss Porta Pro have both reported low ambient sound isolation, but more user tests should be done before these can be recommended for these tasks. Several of the participants in the Sound Guide case reported dislike with headphones with plugs (in-ear) so individual preferences needs to be accounted for here.

Wearable

Smartphones are wearable in the extent that they are always carried around with the users. With traditional navigation aids that rely solely on GPS-positioning, it matters little how it is carried around as long as the GPS-sensor manages to make readings of satellites. It can be placed in the pocket of a jacket or in the trousers, in the bag or held in hand while a headset delivers the information needed.

With navigation aids that utilize gyroscope and compass sensors the placement of the device is more influential. Since the sensors that detect orientation is located on the device itself, it has to be in a fixed orientation with respect to the user. The application need to be able to infer this orientation to be able to correctly display heading information to the user. There exist many approaches to meet these requirements.

In the field-tests with the Sound Guide prototype, the device was held flat in hand in an approximately stable orientation with respect to the user. This proved to be a good position for scanning techniques where the user are able to swipe the device back and forth to explore their environment. For applications that primarily deliver 'look around' functions, this may be a suitable method of use. For applications like route guidance where the pedestrian want to use directional guidance throughout the route, such a method of use would be impractical. The participants manage to use the device in this way during the field-tests, but they stressed the fact that they wanted to be able to have their hands free while walking in everyday use. If the device required one hand and the dog required the other, they would struggle to open gates or carry shopping bags.

As mentioned earlier, a possible solution could have been to let pedestrians carry the device in a case with a neck strap that position the smartphone in upright position at the center of their chest. This enable free access to the device, reasonably available hardware to purchase, and it does give a predictable orientation with respect to the user. Proximity sensors can be used to indicate what side of the device that is facing forward, and accelerometer sensors can be used to detect that the device is in the upright position before giving directional information. Several participants from the Sound Guide case reported such a wearable solution to be suitable for their use. Other studies have also showed that visually impaired people report that a wearable navigation aid would be acceptable [**?**]. This would also comply with the guidelines of **?**, where the GPS aid should not interfere with the primary mobility aid as the cane or the dog.

With respect to interfaces that require scanning behaviour, body orientation detection maybe not be the best choice. **?** participants reported it to be difficult to use the haptic pointer interface when orientation was locked to the body. No such responses were given with respect to the virtual mode, where spatial audio was used.

We experienced *gimbal lock* problems in detecting rotation while the iPhone was held upright. This reported yaw value flicked back and forth a 180° if the pitch angle crossed the max value, and resulted in the direction information being unusable in upright position. Gimbal lock is an inbound limitation of *euler angles*[7] where two axis overlap and make representation of orientation ambiguous. We did not manage to avoid this problem in development and had to reside on using the prototype in a lying position. A possible solution could be to use quaternion to represent orientation, but this was not tested during development due to its large implications on the methods developed.

The neck strap solution require the user to be aware of and always follow the requirement of how the devices needs to be carried to function properly. This may be a sufficient solution, but it could be possible to design applications that do not pose these requirements on the user. If special hardware can be developed, the orientation sensors may be positioned on the headset, locked to the head orientation. The smartphone could then be placed in the pocket without further restrictions. A more elegant solution could be to use GPS-positioning data to detect the orientation of the device in the pocket with respect to the direction of locomotion. If the device's GPS-sensors can see that the user is moving north and its compass sensor show that the device is facing east, it may infer that its own orientation is 90° of the users orientation and deliver direction information with respect to that offset. How stable this solution would have been in practical use needs to be seen in testing.

6.2.6 Independence of Movement

The need for assistance of other people in navigation infer on the fundamental need of independence of movement. Visually impaired people should have the possibility to travel safely, move freely in the environment and not be dependent on others to help them [**?**]. Electronic M&O aids has the potential to compensate for the lack of sight in navigation, and help visually impaired pedestrians to navigate freely in the environment.

The USCB Personal Guidance System was tested on sighted persons to compare how the visually impaired participants compared in speed and distance walked. They found that the average visually impaired pedestrian used about twice as much time as sighted pedestrian using the same system [**?**]. If the sighted pedestrians were blindfolded to make the conditions more equal, the blindfolded pedestrians without M&O training used about twice as much time as visually impaired pedestrians that were used to navigate without sight. This show that orientation and mobility without sight is a skill that needs to be learned.

These numbers indicate that M&O aids have a long way to go if the goal is to compensate fully for the lack of sight. However, this may be an

[7] *Euler angles* is used to describe the orientation of an object based on three axis of orientation.

incorrect goal to start with. If visually impaired persons through new tools are given the freedom to navigate freely, safe and with dignity - a little less speed should be given less priority.

This does not only require electronic M&O aids that let users be able to travel from A to B. The control should be given the users, that may act freely and use these devices in the ways they found useful. As **?** proclaims: *"automation is dangerous when it takes too much control from the user"* [p 197]. The Sound Guide prototype only offered predefined routes and helped the participant to traverse these routes from start to finish. However, this concept does not make the user truly independent. They still need assistance in finding and creating the routes. If not other information is presented, the application will guide the users on fixed routes and will not let them move freely in the environment beyond the recorded routes.

Such development needs to be taken in steps, but care should be taken to design aids that give as much control to the user as possible. **?** requested that visually impaired users *"should be allowed to interact freely and directly with every element in the virtual environment. Even tough sometimes some type of guidance is necessary; users should be free to decide whether or not to accept the suggestion made"*[p 384]. The design of the Sound Guide prototype could have needed a more open approach to accommodate these requirements.

The request for giving control to the users is not just on political or moral grounds. An application is not always present, or it could malfunction during use. The users should be given possibilities to validate the information they get from the application through the perception of ambient sounds and by other M&O aids. It may be times where the user need to take control of the navigation or validate the directions given by the aid. One of the visually impaired participants of the Sound Guide case once found himself on the highway because his GPS had guided him to a new route that utilized the highway even though it was not safe for pedestrians. Technology is never fail proof and the user should therefore have the ability to validate the instructions given through other means. It may also be times when the battery of the M&O aid dies before the pedestrian reach his/her target. The device should then have given enough information about the environment that the user is able to know where he/she is and be able to get back through other means.

6.2.7 Summary

The informal field test conducted, and the low number of participants, have made it hard to make conclusive findings in the Sound Guide case. The prototype developed was met with positive feedback from the participants, and showed much potential as a navigation aid for visually impaired pedestrians. The impact of using spatial audio in this prototype was not easy to determine, due to the use of heading confirmation.

A number of design challenges was revealed, where both technical limitations and interface aspects were found to produce breakdowns in the design. A new method of presenting the coming turn was proposed, together with a way of resolving front/back confusion. How the device should be carried by the user, was also discussed.

Feedback from users and relevant theory have exposed user needs that are not currently met by the prototype. The users want to be able to control when they receive information from the device. They should also be given the possibility to navigate freely without being locked by prerecorded routes.

The use of smartphones as a hardware platform proved promising, both from a developers point of view and the users. The ability to do real-time filtering of spatial audio on proprietary smartphone platforms is still very limited.

Chapter 7

Conclusion

In this work, we have developed a prototype of an application for an open smartphone platform that utilize spatial audio for directional guidance in route navigation for visually impaired pedestrians. Theory from a large range of fields was needed to accomplish this. Limitations of the human perception and supported technology was here two of the main conditions.

The prototype was tested using both qualitative and quantitative methods on a small sample of visual impaired participants. The qualitative results were analysed with content analysis and formed the foundation of the evaluations of the prototype developed and its concept of use.

The theoretical and practical work conducted in this study also resulted in several interaction design guidelines for the use of spatial audio in mobile applications.

7.1 Final Results

The human perception of sound localization puts many limitations on the use of spatial audio in interaction design. To utilize its abilities, it is vital to understand these limitations and be guided by them in the design of spatial audio interfaces. The guidelines inferred for the first research question in this study act as a good starting point for interaction designers that want to use spatial audio techniques in their projects.

The guidelines talk about the horizontal front arc as the optimum area for sound localization. Elevation changes and the rear side is possible to use, but should be approached with care. Interactive displays that use head tracking or body tracking are highly beneficial. Such interaction reduce localization problems, makes spatial audio more realistic and help in selling the mental model to users. Poor quality headphones can be used, but the more complex sound localizations the higher quality headphones are required. In a mobile context one has to look at ambient sounds as something that can influence the display, and something that can be influenced by the display. The psychoacoustical cues in sound localization should not be seen as strict requirements that have to be included, but rather as a toolbox that can be used in creative interface design.

Later research may disproof some of these guidelines or new commercial products might prove them wrong. This is neither a complete list, where several other guidelines are needed to complete the picture. We do still hope that these guidelines may contribute to the field and help others in developing new and exciting interaction concepts based on spatial audio.

The second research question put spatial audio on a test where its use in an orientation aid for visually impaired users was tried out. The development and concept building produced many compelling findings that ranged from hardware support, sound filtering, sensor use and choice of headphones.

Even though the prototype suffered from technical instabilities during the field tests, the responses were positive. The blind participants saw potential in this technology and how it could be used in providing directional information. The explicit utilization of spatial audio did seem to vary in the field test due to the introduction of heading confirmation. The use of heading confirmation is evaluated to be advisable, but other research indicate that it is not necessarily vital in directional guidance.

The use of spatial audio on open smartphone platforms show potential, and we expect that more real-time support for spatial audio will come in the near future. Spatial audio has showed that it can be used to produce richer and more intuitive auditory interfaces. Together with smartphones and sensor technology, we believe spatial audio can be used to give visually impaired people easier access to their environment and help them be more independent in their everyday life.

7.2 Future Work

This work has barely touched the surface when it comes to the use of spatial audio in mobile applications. Much work can be done in refining the guidelines formulated in this thesis. Since the guidelines are mostly inferred from theory, it would be beneficial to device experiments that put them to the test to see their practical appliance in auditory displays. More theory can also be approached for triangulation.

More work is also needed in developing a route guidance application for visually impaired pedestrians that use spatial audio for directional guidance. Formal user tests with larger samples of participants could be beneficial to learn more about the different aspects of these applications. Use concept, sound design, localization accuracy, methods of interaction, hardware implementation, contextual influence and user acceptance are all topics that are highly relevant to study further.

The prototypes developed in this work could be used to do more thorough testing of sound localization than what was achieved. It would be interesting to compare spatial audio with stereo panning, different general HRTF profiles against each other, or the effects of applying acoustical cues like reverberation or room reflections on the ability to localize sound sources. Different headsets could also be compared using the same setup to see how much they influence this ability.

It could also be valuable to do some Wizard of Oz testing, where the directions are controlled by the researcher, to avoid the instability of the location and orientation sensors on the device. Logging of the actual path walked could also be of interest in identifying how well route directions are communicated, and should be easy to implement in the prototype.

Spatial audio displays can be used in other tasks than pedestrian navigation. During our work with this project, we have discovered many interesting and related topics that could have been studied further: Can spatial audio produce more effective and information rich menu systems in screen readers where multiple elements can be presented simultaneous? Could camera or proximity sensors be used to detect obstacles in the close environment around the user and presented their location through spatial audio? Can spatial audio be used as a method of communicating directional information between a supervising control room and a field worker operating with a head mounted camera?

Many more fields of used can be proposed for spatial audio. It is now up to the research community or commercial developers to discover the most promising of those. We believe spatial audio may become highly used in auditory displays of the future and hope our work can contribute in this process.

Appendix A : Sound Guide Prototype

The Sound Guide prototype project is freely distributed under the Creative Commons licence Attribution-NonCommercial-ShareAlike 3.0 Unported [1] and can be downloaded from:

`https://github.com/joakimbording/SoundGuide`

[1] `http://creativecommons.org/licenses/by-nc-sa/3.0/`

Appendix B : Spatial Audio Tester Prototype

The prototype developed to test the accuracy of spatial audio perception is freely distributed under the Creative Commons licence Attribution-NonCommercial-ShareAlike 3.0 Unported [2] and can be downloaded from:

`https://github.com/joakimbording/SpatialAudioTester`

[2] `http://creativecommons.org/licenses/by-nc-sa/3.0/`

Appendix C : Interview Guide

The interview guide written before the interviews and used during the interviews by the researcher. This interview guide was sent to the Norwegian Social Science Data Services (NSD).

The native language of the participants was Norwegian so the interview guide are written and presented in this language.

Intervju Guide 1

For Masteroppgaven:
"Bruk av romlig lyd i mobil navigasjons applikasjon for synshemmede" ved Joakim Bording 2012
Versjon 1. Oppdatert 29.11.2011

Formål:
Formålet med prosjektet er å se på tekniske og praktiske utfordringer og muligheter knyttet til bruk av romlig lyd i mobil navigasjon for synshemmede. Hvordan kan dette gjøres på dagens mobile platformer? Hvor nøyaktig er oppfattelsen av posisjon? Kan disse teknikkene hjelpe blinde i å navigere seg utendørs? Målet er å komme et steg nærmere en applikasjon som kan tilbys sluttbrukere.

Intervju og test lokasjon:
I lokalene til MediaLT i Jerikoveien 22 i Oslo. Første del finner sted i et lukket møterom, andre del finner sted utendørs i lokalmiljøet rundt bygget.

Deltaker profil:
En voksne med eller uten synshemmninger som er vant med å bruke mobiltelefon i hverdagen. Han/hun har ikke møtt intervjuholder før, men har trolig vært med på lignende tester tidligere og kjenner trolig omgivelsene intervjuet blir holdt i.

Overordnet plan for testen:
Del 1: Introduksjon og informasjon om testen. Deltakeren får høretelefoner og en mobil som kjører en prototype applikasjon som simulerer posisjonering av lydkilder rundt omkring i rommet. Oppgaven til deltakeren er å lokalisere disse, rotere kroppen mot de og trykke på skjermen når han/hun anser lydkilden som rett foran. Øvelsene blir avsluttet med intervju.

Del 2: Testsleder og deltaker beveger seg ut. Deltaker blir gitt høretelefoner og en mobil som kjører en utvidet prototype som guider brukeren langs oppsatte ruter. Deltakeren blir bedt om å navigere seg fram til en eller flere av målene ved å bruke prototypen. Deltakeren blir bedt om å tenke høyt under testen og kan også få direkte spørsmål. Etter øvelsen blir det et kort intervju før testen blir avsluttet.

Guide Del 1:

Introduksjon

1. Ønsk velkommen - Te/Kaffe?
2. Introduser megselv. Student. Master i design, bruk og interaksjon. Individuell masteroppgave.
3. Romlig lyd, eller 3D lyd, er lyd som blir oppfattet å komme fra en bestemt retning. Knips. Dette kan gjenskapes tildels via øretelefoner. Formålet med prosjektet er å se på hvordan bruk av slik romlig lyd kan støtte synshemmede i navigering. Jeg tar utgangspunkt i hva som kan gjøres på dagens mobiler, så i dag vil vi prøve prototyper på en iPhone 4S med standard øreplugger. Jeg vil se på de tekniske begrensningene og ønsker å teste hvordan forskjellige lydoppsett oppleves.
4. Fortelle kort om planen for dagen. Opptak av intervju
5. Opplys om deltakerens rettigheter som deltaker i forskningstudie. Gå igjennom punktene i samtykkeskjema og be deltakeren signere hvis deltakeren har forstått og godtar innholdet.

Forberedelser

1. Be deltakeren plassere seg midt i rommet. Monter øretelefoner.
2. Test at høyre og venstre blir oppfattet like høye med et lydklipp.
3. Prøv å hold mobilen flat og pekende ut fra kroppen slik.
4. Forklar oppgaven som kommer. "*Du vil bli tatt tiden på og nøyaktighet på peking blir notert. Du vil kunne høre dine resultater etterpå hvis du vil. Men ikke stress og ikke vær pinlig nøyaktig på sikte. Utfør hver oppgave smidig og jevnt. Bare å spør hvis du har spørsmål.*"

Test

1. Tre runder med litt forskjellig oppsett. Deltakeren får en kort instruksjon før hver runde, mens hver runde blir utført i stillhet.
2. Ti forsøk med et tilfeldig plassert lydikon som gir lyd kontinuerlig.
3. Ti forsøk med et tilfeldig plassert lydikon som avgir kun en kort lyd
4. Ti forsøk med flere tilfeldige plasserte lydikon hvor de skal lokalisere posisjonen til det lydikonet de kjenner fra før.

Intervju (semistrukturert):

Opplys om at lydopptakeren blir startet.

1. *"Hvordan opplevde du disse oppgavene?"*
2. *"Var det runder du opplevde som mer utfordrene? Noen mer lette?"*
3. *"Var du sikker på posisjonen når du trykket på knappen?"*
4. *"Bruker du mobilen aktivt i hverdagen?"*
5. *"I hvilke situasjoner bruker du øretelefoner?"*

6. *"Pleier du å bruke noe teknologi til navigering utendørs?"*
7. *"Hvordan navigerer du deg på nye områder?"*
8. *"Hva savner du av muligheter i dagens hjelpemidler i forhold til navigasjon?"*
9. *"Hva slags type informasjon anser du som det viktigste i en navigasjonsapplikasjon"*
10. *"I hvilke situasjoner vil du kunne ha behov for en navigasjonsapp som tar i bruk lyd"*

Guide Del 2:

Introduksjon

1. Forklar oppgaven som kommer og dens formål. *"Vil vare ca 20 min. Jeg vil følge deg hele veien og vil kunne assistere deg i den grad du ønsker det. Du blir ikke tatt tiden på, det jeg er interessert i er hvordan du opplever å bruke dette som et verktøy. (Beskriv ruten) Snakk høyt over de tanker du gjør deg rundt navigering og bruk og tolkning av applikasjonen"*
2. *"Er du komfortabel med å navigere deg utendørs her på dette viset?"*

Forberedelser

1. Ytterklær og transport ut.
2. Monter øretelefoner. Test at høyre og venstre er riktig
3. Opplys om at lydopptakeren blir startet.
4. Roter brukeren i 360 grader og angi hvilken retning lyden indikerer

Test

1. Finn veien ved hjelp av en lagret rute
2. Finn veien tilbake ved hjelp av samme rute

Mulige intervjuspørsmål under testen (ustrukturert)

1. *Observer behov for headtracking*
2. *"Tenk gjerne høyt når du stusser på ting"*
3. *"Hvorfor valgte du å gå den veien?"*
4. *"Hva ønsker du å få informasjon om fra en slik applikasjon?"*
5. *"Kunne lydikonene blitt presentert på en bedre måte?"*
6. *"Hvordan føles det å navigere seg basert på posisjon av lyd når du går ute?"*
7. *"Ville du kunne ønsket å brukt en slik applikasjon i dagliglivet?"*
8. *"Er du komfortabel med å gå med øreklokker som disse ute?"*
9. *"Hvis du skulle navigert denne strekningen normalt, hvordan ville du da gjort det?"*
10. *"Hvordan finner du butikker og lokaler til vanlig?"*

Avslutning

1. *"Er det noen ting du ønsker å få frem som vi ikke har pratet om?"*
2. Opplys om at lydopptakeren blir skrudd av.
3. Takke stort for deltakelsen i studiet!
4. Gi kontaktdata med kontaktopplysninger til oss hvis spørsmål

Appendix D : Consent Form

The consent form used in the study conducted in this thesis. This consent form was sent to the Norwegian Social Science Data Services (NSD). Each participant was asked to sign this paper. Phone numbers are anonymized.

The native language of the participants was Norwegian so the consent form are written and presented in this language.

Forespørsel om å delta på intervju og test i forbindelse med en masteroppgave

Jeg er masterstudent i informatkk: design, bruk og interaksjon ved Universitetet i Oslo og holder nå på med den avsluttende masteroppgaven. Temaet for oppgaven er bruk av romlig lyd i mobil navigasjon applikasjon for synshemmede. Romlig lyd kan også kalles 3D lyd og omfatter lydkilder som oppfattes som posisjonert i rommet i forhold til brukeren. Jeg har utviklet noen prototyper som tar i bruk romlig lyd og ønsker å gjøre brukertester for å avdekke utfordringer og muligheter knyttet til teknologien.

For å finne ut av dette, ønsker jeg å intervjue og gjøre brukertester på 10 voksne som er vant til å bruke mobiltelefoner i hverdagen, hvor noen har en synshemmning som gjør at de ikke kan se. Oppgavene vil dreie seg om å teste en mobil applikasjon ved å lokalisere retningen forskjellige lyder kommer fra og prate om erfaringer knyttet til dette. Dette blir gjort for å avgjøre hvor nøyaktig teknologien bak romlig lyd er og dermed hvor anvendelig den er til forskjellige formål.

Det vil også bli gjort en test hvor deltakerne kan navigere seg utendørs ved hjelp av en applikasjon og prate om erfaringer knyttet til dette og navigering i hverdagen generelt. Formålet med dette er å avdekke begrensninger og muligheter ved anvendelse av romlig lyd i mobil navigasjon applikasjon for synshemmede. Studiens formål er å teste teknologiens anvendelighet og ikke deltakeren som brukere direkte.

Jeg vil bruke båndopptaker og ta notater mens vi snakker sammen. Jeg vil observere og ta notater under oppgavene. Hele testen vil ta omtrent en time, og vi blir sammen enige om tid og sted.

Det er frivillig å være med og du har mulighet til å trekke deg når som helst underveis, uten å måtte begrunne dette nærmere. Dersom du trekker deg vil alle innsamlede data om deg bli anonymisert. MediaLT vil formidler kontakten med mulige deltakere, hvor undertegnede student ikke kjenner deltakernes identitet før de samtykker til å delta på testen. Opplysningene vil bli behandlet konfidensielt, og ingen enkeltpersoner vil kunne gjenkjennes i den ferdige oppgaven. Opplysningene anonymiseres og opptakene slettes når oppgaven er ferdig, innen utgangen av 2012.

Hvis det er noe du lurer på kan du ringe meg på ## ## ## ## eller sende en e-post til joakim@bording.no. Du kan også kontakte min veileder Jo Herstad på Instituttet for Informatikk på Universitetet i Oslo på telefonnummer ## ## ## ##, eller min veileder Morten Tollefsen på MediaLT på telefonnummer ## ## ## ##.

Studien er meldt til Personvernombudet for forskning, Norsk samfunnsvitenskapelig datatjeneste A/S.

Med vennlig hilsen
Joakim Bording
####
Oslo

Samtykkeerklæring:
Jeg har mottatt informasjon om studien i bruk av romlig lyd i mobil navigasjon applikasjon for synshemmede og ønsker å delta på testen.

Signatur …………………………………   Telefonnummer ……………………………..

Appendix E : Sound Localization Accuracy Test Results

The results from the test that are presented in section 5.1 is here presented. Data that was judged as not valid is not included. The table has the following content:

1. *USER*: The participant.
2. *ROUND*: Each setup type was administered as one round.
3. *POST:OFFSET*: The angular offset of the reported direction with respect to the sound icon.
4. *REL:OFFSET*: Same as post.offset only with positive values. Served as the foundation for calculations.
5. *PRE:OFFSET*: Position of sound icon with respect to the user when first presented.
6. *TIME.USED*: Number of seconds the participant used on the task.
7. *SETUP.TYPE*: The type of sound administered.

USER	ROUND	AIM	POST.OFFSET	**REL.OFFSET**	PRE.OFFSET	TIME.USED	SETUPTYPE
1	1	6	-3,59	**3,59**	12,53	2,56	ApplePurrPanoramaLoop
1	1	7	-130,36	**130,36**	197,16	4,41	ApplePurrPanoramaLoop
1	1	8	-46,20	**46,20**	122,54	1,66	ApplePurrPanoramaLoop
1	1	9	7,60	**7,60**	298,51	17,41	ApplePurrPanoramaLoop
1	1	10	37,30	**37,30**	118,92	1,71	ApplePurrPanoramaLoop
1	1	11	-6,48	**6,48**	345,32	6,22	ApplePurrPanoramaLoop
1	1	12	15,05	**15,05**	274,25	17,49	ApplePurrPanoramaLoop
1	1	13	-31,94	**31,94**	54,36	4,38	ApplePurrPanoramaLoop
1	2	4	-28,38	**28,38**	36,84	5,11	ApplePurrPanningLoop
1	2	5	1,56	**1,56**	356,71	4,14	ApplePurrPanningLoop
1	2	6	-17,68	**17,68**	302,57	7,59	ApplePurrPanningLoop
1	2	7	4,35	**4,35**	135,42	2,13	ApplePurrPanningLoop
1	2	8	-18,98	**18,98**	316,90	1,64	ApplePurrPanningLoop
1	2	9	-36,09	**36,09**	23,49	7,07	ApplePurrPanningLoop
1	2	10	-44,07	**44,07**	316,68	2,28	ApplePurrPanningLoop
1	2	11	-34,89	**34,89**	115,71	5,64	ApplePurrPanningLoop
1	2	12	-26,99	**26,99**	152,44	4,48	ApplePurrPanningLoop
1	2	13	11,27	**11,27**	347,74	9,64	ApplePurrPanningLoop
1	3	5	26,60	**26,60**	344,53	2,00	ApplePurrPanning
1	3	6	51,02	**51,02**	157,86	4,00	ApplePurrPanning
1	3	7	-31,20	**31,20**	3,74	3,07	ApplePurrPanning
1	3	8	24,39	**24,39**	220,44	3,32	ApplePurrPanning
1	3	9	96,00	**96,00**	157,03	1,79	ApplePurrPanning
1	3	10	-55,76	**55,76**	28,68	3,49	ApplePurrPanning
1	3	11	-31,55	**31,55**	196,28	3,46	ApplePurrPanning
1	3	12	-168,69	**168,69**	220,43	2,48	ApplePurrPanning
1	3	13	-52,22	**52,22**	158,71	1,25	ApplePurrPanning
1	4	4	-10,82	**10,82**	44,68	1,31	ApplePurrPanorama
1	4	5	-41,51	**41,51**	53,88	2,21	ApplePurrPanorama
1	4	6	53,03	**53,03**	209,48	3,12	ApplePurrPanorama
1	4	7	-70,40	**70,40**	227,90	1,86	ApplePurrPanorama
1	4	8	11,02	**11,02**	245,28	3,09	ApplePurrPanorama
1	4	9	-2,66	**2,66**	107,71	4,14	ApplePurrPanorama
1	4	10	-54,18	**54,18**	246,29	1,88	ApplePurrPanorama
1	4	11	-41,72	**41,72**	348,25	4,47	ApplePurrPanorama
1	4	12	-6,43	**6,43**	19,19	2,77	ApplePurrPanorama
1	4	13	-16,03	**16,03**	343,33	3,81	ApplePurrPanorama
1	5	4	141,91	**141,91**	132,81	9,57	ToneFrequencyLoop
1	5	5	-28,70	**28,70**	260,12	18,26	ToneFrequencyLoop
1	5	6	-49,17	**49,17**	40,29	6,66	ToneFrequencyLoop
1	5	7	42,20	**42,20**	198,49	6,83	ToneFrequencyLoop
1	5	8	66,56	**66,56**	38,31	5,10	ToneFrequencyLoop
1	5	9	-99,91	**99,91**	127,05	4,68	ToneFrequencyLoop
1	5	10	24,46	**24,46**	355,13	4,63	ToneFrequencyLoop
1	5	11	51,58	**51,58**	63,14	7,48	ToneFrequencyLoop
1	5	12	-39,42	**39,42**	270,41	6,11	ToneFrequencyLoop
1	5	13	31,00	**31,00**	204,86	6,40	ToneFrequencyLoop
3	1	4	-29,36	**29,36**	170,39	11,77	ApplePurrPanoramaLoop
3	1	5	-3,02	**3,02**	61,35	18,95	ApplePurrPanoramaLoop

3	1	6	1,20	**1,20**	290,40	24,84	ApplePurrPanoramaLoop
3	1	7	6,81	**6,81**	307,15	7,17	ApplePurrPanoramaLoop
3	1	8	5,19	**5,19**	140,17	5,61	ApplePurrPanoramaLoop
3	1	9	8,77	**8,77**	39,78	6,96	ApplePurrPanoramaLoop
3	1	10	16,38	**16,38**	318,04	3,71	ApplePurrPanoramaLoop
3	1	11	8,38	**8,38**	303,85	3,83	ApplePurrPanoramaLoop
3	1	12	19,33	**19,33**	216,89	4,04	ApplePurrPanoramaLoop
3	1	13	20,62	**20,62**	218,25	4,09	ApplePurrPanoramaLoop
3	2	4	23,97	**23,97**	294,17	3,86	ApplePurrPanningLoop
3	2	5	150,13	**150,13**	154,60	3,90	ApplePurrPanningLoop
3	2	6	-0,57	**0,57**	223,75	6,13	ApplePurrPanningLoop
3	2	7	-16,96	**16,96**	56,67	5,35	ApplePurrPanningLoop
3	2	8	16,96	**16,96**	264,20	4,14	ApplePurrPanningLoop
3	2	9	-5,79	**5,79**	335,87	4,33	ApplePurrPanningLoop
3	2	10	-3,06	**3,06**	254,31	7,20	ApplePurrPanningLoop
3	2	11	2,00	**2,00**	322,43	5,96	ApplePurrPanningLoop
3	2	12	-35,52	**35,52**	160,29	7,30	ApplePurrPanningLoop
3	2	13	-16,61	**16,61**	3,93	8,43	ApplePurrPanningLoop
3	3	4	3,95	**3,95**	109,48	1,45	ApplePurrPanning
3	3	5	-26,85	**26,85**	130,94	2,19	ApplePurrPanning
3	3	6	-85,00	**85,00**	263,99	2,67	ApplePurrPanning
3	3	7	-38,41	**38,41**	202,77	1,30	ApplePurrPanning
3	3	8	103,91	**103,91**	308,62	1,40	ApplePurrPanning
3	3	9	-157,80	**157,80**	157,42	3,52	ApplePurrPanning
3	3	10	-45,29	**45,29**	352,18	2,32	ApplePurrPanning
3	3	11	150,89	**150,89**	67,62	1,87	ApplePurrPanning
3	3	12	-117,96	**117,96**	81,47	1,73	ApplePurrPanning
3	3	13	67,65	**67,65**	125,63	1,40	ApplePurrPanning
3	4	4	128,52	**128,52**	29,01	2,13	ApplePurrPanorama
3	4	5	20,14	**20,14**	297,67	1,58	ApplePurrPanorama
3	4	6	-83,79	**83,79**	69,36	1,45	ApplePurrPanorama
3	4	7	24,65	**24,65**	82,47	1,40	ApplePurrPanorama
3	4	8	-8,03	**8,03**	2,74	0,97	ApplePurrPanorama
3	4	9	-36,71	**36,71**	84,11	1,97	ApplePurrPanorama
3	4	10	25,32	**25,32**	76,77	1,37	ApplePurrPanorama
3	4	11	-21,45	**21,45**	239,75	1,34	ApplePurrPanorama
3	4	12	7,05	**7,05**	342,27	1,59	ApplePurrPanorama
3	4	13	-49,89	**49,89**	225,58	1,23	ApplePurrPanorama
3	5	4	-5,02	**5,02**	117,73	6,97	ToneFrequencyLoop
3	5	5	2,17	**2,17**	188,87	6,49	ToneFrequencyLoop
3	5	6	6,67	**6,67**	146,06	6,85	ToneFrequencyLoop
3	5	7	6,49	**6,49**	299,72	6,76	ToneFrequencyLoop
3	5	8	4,16	**4,16**	49,57	8,07	ToneFrequencyLoop
3	5	9	12,05	**12,05**	67,60	4,63	ToneFrequencyLoop
3	5	10	7,82	**7,82**	245,27	9,19	ToneFrequencyLoop
3	5	11	-3,61	**3,61**	322,81	3,64	ToneFrequencyLoop
3	5	12	-5,80	**5,80**	85,22	7,26	ToneFrequencyLoop
3	5	13	12,22	**12,22**	150,44	6,27	ToneFrequencyLoop