



2012-12-01

# Asynchronous Ultrasonic Trilateration for Indoor Positioning of Mobile Phones

Viacheslav Filonenko (Thesis)

*Dublin Institute of Technology*

Follow this and additional works at: <http://arrow.dit.ie/dmcthes>

 Part of the [Technology and Innovation Commons](#)

## Recommended Citation

Filonenko,V.: Asynchronous Ultrasonic Trilateration for Indoor Positioning of Mobile Phones, Doctoral Thesis, Dublin Institute of Technology, 2012.

This Theses, Ph.D is brought to you for free and open access by the Digital Media Centre at ARROW@DIT. It has been accepted for inclusion in Theses by an authorized administrator of ARROW@DIT. For more information, please contact [yvonne.desmond@dit.ie](mailto:yvonne.desmond@dit.ie), [arrow.admin@dit.ie](mailto:arrow.admin@dit.ie).



This work is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 License





# **Asynchronous Ultrasonic Trilateration for Indoor Positioning of Mobile Phones**

Viacheslav Filonenko

B.Sc. (Computer Science) Dublin Institute of Technology, 2008

This Thesis is submitted to

the Dublin Institute of Technology

School of Media

for the Degree of

Doctor of Philosophy

December 2012

Research Supervisors:

Dr. James D. Carswell, Head of Spatial Technologies Research, Digital Media Centre, DIT

Dr. Charlie Cullen, Principal Investigator, Digital Media Centre, DIT

Dr. Michela Bertolotto, Senior Lecturer, School of Computer Science and Informatics, UCD

I certify that this thesis which I now submit for examination for the award of Doctor of Philosophy, is entirely my own work and has not been taken from the work of others, save and to the extent that such work has been cited and acknowledged within the text of my work.

This thesis was prepared according to the regulations for postgraduate study by research of the Dublin Institute of Technology and has not been submitted in whole or in part for another award in any other third level institution. The work reported on in this thesis conforms to the principles and requirements of the DIT's guidelines for ethics in research.

DIT has permission to keep, lend or copy this thesis in whole or in part, on condition that any such use of the material of the thesis be duly acknowledged.

Signature \_\_\_\_\_

Date \_\_\_\_\_

# **ABSTRACT**

Spatial awareness is fast becoming the key feature on today's mobile devices. While accurate outdoor navigation has been widely available for some time through Global Positioning Systems (GPS), accurate indoor positioning is still largely an unsolved problem. One major reason for this is that GPS and other Global Navigation Satellite Systems (GNSS) systems offer accuracy of a scale far different to that required for effective indoor navigation. Indoor positioning is also hindered by poor GPS signal quality, a major issue when developing dedicated indoor locationing systems. In addition, many indoor systems use specialized hardware to calculate accurate device position, as readily available wireless protocols have so far not delivered sufficient levels of accuracy. This research aims to investigate how the mobile phone's innate ability to produce sound (notably ultrasound) can be utilised to deliver more accurate indoor positioning than current methods. Experimental work covers limitations of mobile phone speakers in regard to generation of high frequencies, propagation patterns of ultrasound and their impact on maximum range, and asynchronous trilateration. This is followed by accuracy and reliability tests of an ultrasound positioning system prototype.

This thesis proposes a new method of positioning a mobile phone indoors with accuracy substantially better than other contemporary positioning systems available on off-the-shelf mobile devices. Given that smartphones can be programmed to correctly estimate direction, this research outlines a potentially significant advance towards a practical platform for indoor Location Based Services. Also a novel asynchronous trilateration algorithm is proposed that eliminates the need for synchronisation between the mobile device and the positioning infrastructure.

## PUBLICATIONS

Filonenko, V. and Carswell, J., (2009)

"*Hybrid indoor positioning and directional querying on a ubiquitous mobile device*"; in Proceedings of *6<sup>th</sup> International Symposium on LBS & Telecartography*; September 2-4; University of Nottingham, UK; Paper 5.

<http://arrow.dit.ie/dmccon/5>

Filonenko, V. and Carswell, J., (2009)

"*Tracker: indoor positioning for the LOK8 project*"; in Proceedings of *9th. IT & T Conference*; October 22-23; Dublin Institute of Technology, Ireland; Paper 25.

<http://arrow.dit.ie/ittapn/25>

Filonenko, V., Cullen, C., and Carswell, J., (2010):

"*Investigating Ultrasonic Positioning on Mobile Phones*"; in Proceedings of *International Conference on Indoor Positioning and Indoor Navigation (IPIN 2010)*; September 15 – 17; ETH Zurich, Switzerland; Pages 419-426; IEEE Xplore 2010.

[http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=5648235&tag=1](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5648235&tag=1)

Filonenko, V., Cullen, C., and Carswell, J., (2012):

"*Asynchronous Ultrasonic Trilateration for Indoor Positioning of Mobile Phones*"; in Proceedings of *Web and Wireless Geographical Information Systems (W2GIS 2012)*; April 12 – 13; Naples, Italy; Pages 33-46; Volume 7236 of *Lecture Notes in Computer Science*; Springer-Verlag, 2012.

[http://rd.springer.com/chapter/10.1007/978-3-642-29247-7\\_4](http://rd.springer.com/chapter/10.1007/978-3-642-29247-7_4)

Full publications are provided at the end of this document.

# TABLE OF CONTENTS

ABSTRACT .....	ii
PUBLICATIONS .....	iii
TABLE OF CONTENTS .....	v
LIST OF FIGURES .....	vi
LIST OF TABLES .....	viii
1. INTRODUCTION .....	1
1.1. Problem Statement .....	3
1.2. Motivation of the Thesis .....	6
1.3. Aims of the Thesis .....	9
1.4. Document Structure .....	11
2. RELATED RESEARCH AND BACKGROUND .....	13
2.1. Positioning Methods .....	14
2.1.1. Satellite Navigation Systems .....	14
2.1.2. GSM .....	21
2.1.3. Wi-Fi (802.11) .....	28
2.1.4. Bluetooth .....	34
2.1.5. Sound .....	40
2.1.6. Dead Reckoning .....	50
2.1.7. Computer Vision Approach .....	56
2.1.8. Discussion .....	63
2.2. Services .....	66
2.3. Interfaces .....	70
2.4. Summary .....	74
3. PRELIMINARY EXPERIMENTS .....	80
3.1. Ultrasound Generation .....	80
3.2. Signal Design .....	89
3.3. Range .....	96
4. ASYNCHRONOUS TRILATERATION .....	101
4.1. Least Squares Method for TOA Trilateration .....	101
4.2. Novel Least Squares Method for TDOA Trilateration .....	108
4.3. Working Example of TDOA Trilateration .....	117
5. EVALUATION .....	122
5.1. Positioning Accuracy .....	122
5.2. Angle Variation .....	133
5.3. Direction and Signal Obstruction .....	141
5.4. Background Noise .....	147
6. CONCLUSIONS .....	152
6.1. Summary of Work .....	152
6.2. Contributions of the Thesis .....	157
6.3. Future Work .....	160
6.3.1. Directional querying .....	160
6.3.2. Support Multiple Users .....	163
6.3.3. Custom beacons .....	165
6.3.4. Signal Reception Model .....	166
6.4. Overall Conclusions .....	169
7. REFERENCES .....	171
APPENDIX 1. SPECTROGRAMS .....	176
APPENDIX 2. RANGE MEASUREMENTS .....	197
APPENDIX 3. TDOA SOURCE CODE .....	198
APPENDIX 4. ACCURACY EXPERIMENT VALUES .....	200
APPENDIX 5. EQUIPMENT .....	202

# LIST OF FIGURES

Figure 1: 2D position determination with 3 satellites and corrected clock error.....	17
Figure 2: Nearest Access Point .....	23
Figure 3: Trilateration .....	24
Figure 4: RF Fingerprinting .....	25
Figure 5: Relationship between RSSI and Received power .....	38
Figure 6: BeepBeep signal exchange .....	42
Figure 7: A polar plot depicting 6dB/div measurement for a range of frequencies .....	46
Figure 8: Cardioid and Supercardioid polar patterns.....	47
Figure 9: Roll, Yaw and Pitch axis.....	51
Figure 10: Fiduciary marker.....	57
Figure 11: Retro reflectors captured without and with a flash .....	60
Figure 12: The ultrasound generation experiment setup .....	81
Figure 13: A spectrogram of the file played back by all tested smartphones .....	82
Figure 14: Spectrogram for HTC G1 at file volume 80%, device volume maximum – 2.....	84
Figure 15: Spectrogram for iPhone at file volume 60%, device volume maximum.....	85
Figure 16: Spectrogram for HTC Hero at file volume 100%, device volume maximum.....	86
Figure 17: Spectrogram for Nokia Navigator at file volume 20%, device volume maximum .....	87
Figure 18: Spectrogram for Nokia Navigator at file volume 100%, device volume maximum - 2.....	88
Figure 19: Waveforms of the same 22kHz signal saved at 44.1 and 96 kHz sampling rate .....	90
Figure 20: Spectrograms of the same 22kHz signal saved at 44.1 and 96 kHz sampling rate .....	91
Figure 21: Common slope shapes. ....	92
Figure 22: Waveform of the optimal single-frequency signal.....	94
Figure 23: Waveform of the signal where the first half is composed using a different frequency .....	95
Figure 24: Waveform of the signal where the entire envelope is composed using a different frequency	95
Figure 25: The division of space around the phone into sectors .....	97
Figure 26: Polar contour plot for ultrasound energy propagation .....	98
Figure 27: Relationship between signal strength and distance for conditions where phone speaker and microphone point at each other. ....	100
Figure 28: Time of Arrival .....	102
Figure 29: Time Difference of Arrival .....	109
Figure 30: TDOA Trilateration experiment with four microphones and six different smartphone positions .....	117
Figure 31: Layout of the room used for trilateration accuracy experiments.....	125
Figure 32: Raw filter output .....	125
Table 5: Difference between MPV and true position $\pm$ standard deviation.....	126
Figure 33: Average, best and worst results for Experiment 1 (2D Trilateration). ....	127
Figure 34: Average, best and worst results for Experiment 2 (3D Trilateration). ....	128
Figure 35: Average, best and worst results for Experiment 3 (3D Trilateration with calibration) .....	129
Figure 36: Direction and distance from true position to MPV in Experiment 1 (2D Trilateration). ....	130

Figure 37: Direction and distance from true position to MPV in Experiment 2 (3D Trilateration) ....	132
Figure 38: Direction and distance from true position to MPV in Experiment 3 (3D Trilateration with calibration) .....	133
Figure 39: The change of distance (in cm) between MPV and correct position as the speaker is rotated from downward to upward orientation .....	137
Figure 40: The scattering of detected positions and various pitch angles .....	139
Figure 41: The change in the average of distances between each detected position and correct position as the speaker is rotated from downward to upward orientation .....	140
Figure 42: Layout of the room used for direction and signal obstruction experiments .....	142
Figure 43: Positions in which the user stands at check point 1 .....	143
Figure 44: Difference between MPV and true position for each check point, user's position type and orientation .....	145
Figure 45: Percentage of failed positioning attempts for each check point, user's position category and orientation .....	146
Figure 46: Noise generated from jiggling a bunch of keys after 21.5 kHz bandpass filter .....	148
Figure 47: Noise generated from slamming shut a metal drawer after 21.5 kHz bandpass filter .....	149
Figure 48: Signal detection in normal room noise conditions .....	149
Figure 49: Screenshot of a program window that displays current buffer contents after filtering .....	156
Figure 50: Screenshot of a program window that displays detected user position .....	157
Figure 51: Model of reception quality .....	168

## LIST OF TABLES

Table 1: Comparison of positioning methods for smartphones.....	63
Table 2: Comparison of indoor positioning implementations .....	65
Table 3: Sample TDOA Trilateration input.....	120
Table 4: Comparison of TDOA output and expected results .....	121
Table 5: Difference between MPV and true position $\pm$ standard deviation.....	126
Table 6: Difference between MPV and true position $\pm$ standard deviation for different pitch angles .	136
Table 7: Percentage of detected signals for different angles .....	136
Table 8: Difference between MPV and true position $\pm$ standard deviation for each check point, user's position category and orientation .....	144

# 1. INTRODUCTION

In this thesis we explore the problem of accurately positioning a commercial off-the-shelf (COTS) smartphone indoors and offer a solution which is an order of magnitude more accurate than contemporary indoor positioning systems currently available for these devices. Alternative positioning methods that can be used to locate mobile phones indoors are identified and evaluated based on recorded accuracy of individual implementations, advantages and disadvantages of each method in general and limitations of utilised hardware. Our choice of ultrasound trilateration as the primary research direction is explained followed by an overview of our positioning method. The proposed method is uniquely characterised by two key features. The first feature is the use of an inaudible frequency band inherent to standard smartphone sound hardware. It was unknown how well standard mobile phone speakers would be able to produce an ultrasound signal and whether it could be reliably detected by standard microphones accurately enough to exploit the signal's time-of-flight for positioning. These questions are investigated through lab experiments. The second feature is a novel asynchronous trilateration algorithm, which makes positioning substantially easier by eliminating the need for clock synchronisation between transmitter and receiver at the expense of one extra control point (microphone). It is shown how the asynchronous trilateration algorithm is derived and how well it performs in theory and in real-world situations. In order to determine strengths and weaknesses of the proposed approach, it was implemented and tested for accuracy and reliability. Poor obstacle penetration and the highly directional nature of ultrasound is also evaluated in various experiments and their impact on the proposed positioning method is discussed.

Positioning has a large variety of potential uses on smartphones, which have become popular in the recent years. Also technological advances of smartphones have changed how we view mobile devices in general. We now want to be able to do all the same things we usually do on a desktop computer, except on the move, e.g. blog, chat, surf, play games, watch TV or listen to Radio. At the same time mobile phones are different in that they have to be light and ideally should fit in a pocket, consequently the screen is relatively small and there is not much interface surface to interact with. This characterizes how user interface design for mobile software needs to be approached and suggests that alternative forms of interaction such as voice and gesture may play an important role. These, unfortunately, are not always acceptable. Due to the mobile nature of the device, a user will often be surrounded by other people and other factors they have no control over, such as noise. Normally a user will find it awkward to wave/shake their phone or raise their voice in public. A better alternative is to reduce, and if possible, remove unnecessary user input. Since the device is mobile, the best way to accomplish this is through spatial awareness.

Spatial awareness is the ability of a system to be aware of its surroundings, either in terms of location or proximity to certain objects that might be of interest to the user. Very soon it may become an essential part of our experience with mobile devices, thanks to a few factors. First of all, fast and relatively cheap internet connection is now available for most mobile phone users. Spatially aware systems need data to work with, and without internet connection the scope for possible applications is limited. Secondly, today's smartphones have built-in sensor hardware that may be used to estimate a user's location in one way or another. GPS receivers for example are a de-facto standard positioning system that many LBS services rely on. Without this sort of sensor hardware, attempting to make a universal, reliable and accurate location based service would be severely limited. And finally, every popular mobile platform already has its

own application store available via internet connection. For example Apple and Google already have their “App Store”<sup>1</sup> and “Google Play”<sup>2</sup>, Nokia has OVI Store<sup>3</sup> and Microsoft has Marketplace<sup>4</sup> for Windows Mobile. Besides generating profits for the corporations with very little effort, the stores attract thousands of developers with the possibility to earn money via microtransactions. There are already such 3<sup>rd</sup> party “Apps” available (e.g. Layar and Wikitude) and with all of the factors mentioned above, spatial awareness will continue to gain in popularity both among developers and users.

There are various ways spatial awareness can be used to make a user’s life a bit easier or more informative. The most basic application would be prioritising search results. For example if the user enters “nearest restaurant” as a web search, the system should give him details of a nearby restaurant rather than a standard search result. A more advanced example would be directional querying, which allows a user to point his phone at some building or monument on the street and learn what the place is. In other words, if a mobile device knew its exact position and orientation at any time, such an ability would become interesting for a myriad of new LBS applications (Liu 2011).

## 1.1. Problem Statement

Indoor and outdoor positioning are normally considered as two separate problems. The reason being a solution designed for one environment either will not work in the other or will require significant modification (Schiller 2004; Kolodziej 2006). Below are

---

<sup>1</sup> "App Store" Accessed 12 November, 2012, at <http://www.apple.com/iphone/from-the-app-store/>

<sup>2</sup> "Google Play" Accessed 12 November, 2012, at <https://play.google.com/store>

<sup>3</sup> "Nokia Store" Accessed 12 November, 2012, at <http://store.ovi.com/>

<sup>4</sup> "Windows Phone" Accessed 12 November, 2012, at <http://www.windowsphone.com/en-us/store>

some differences that need to be taken into consideration when designing or modifying a mobile locationing system:

- It is usually easier to place equipment (e.g. Bluetooth beacons, Wi-Fi routers etc.) around an indoor environment. Densely distributing valuable equipment over a large outdoor area is generally inefficient and problematic.
- Outdoor environments may change unpredictably over time. Most of the time we are talking about smaller objects either changing place, appearing or disappearing. Furthermore, renovation, construction or demolition of whole buildings is also possible. Though the same is true for indoor environments, administrators of a positioning system are a lot more likely to be notified in advance about major changes to the environment structure and possibly provided with accurate information about the changes.
- Generally indoor environments require higher accuracy to be useful for practical LBS purposes. This is because when indoors we are dealing with objects and distances at a smaller scale. While accuracy of +/- 10 meters may be good enough to direct someone to a cafe or a bus stop, indoors it would mean we are not sure in which room the user currently is.
- When used indoors, electromagnetic signals can suffer from fading and multipath propagation when they encounter walls, windows, and other structures. Services that rely on satellite signals such as GPS do not work indoors at all because the satellite signal requires a direct line-of-sight to the receiver.

Today, global navigation systems handle outdoor positioning and navigation fairly well. All of these systems rely on signals broadcast by satellites. Other outdoor systems usually are aimed to cover only densely populated areas and are designed to either

replace GPS or enhance its accuracy. They utilise alternative signal sources such as Wi-Fi or GSM because they are abundant in urban environments while GPS suffers from various problems caused by the density of high buildings around the user. GLONASS, which was originally designed to remove Russian army dependence on American GPS, is mostly used in a commercial environment to backup or correct GPS on the same device. In other words GPS is currently the backbone of outdoor navigation and will most probably remain so for years.

It is a different picture in the domain of indoor positioning. Strikingly there is not a single widely accepted commercial indoor locationing solution for mobile devices (Kolodziej 2006). That could be because most such systems do not function well in a completely new environment as each locationing system has to be carefully adjusted for each new location. For example, a locationing system may need its own set of transmitting beacons distributed around the environment where a more dense distribution will result in higher accuracy. It has been shown that really high accuracy (around 3 cm) can be achieved through the use of ultrasound (Harter 1999; Addlesee 2001; Randell 2001; Priyantha 2005), although so far such results were obtained only with custom built hardware not readily available to general users. Electromagnetic waves can be used as well as they are more accessible in the form of Bluetooth/Wi-Fi, but the accuracy will not be as good (around 3 m) (Thapa 2003). Another completely different approach utilises optical recognition. No beacons are needed but the system does need some known landmarks to recognise and a clear view of them, so it is more effective in a static environment rather than a dynamic one with lots of people moving around.

## 1.2. Motivation of the Thesis

Taking into consideration the difficulties associated with current indoor locationing systems, we believe that developing a solution that works on off-the-shelf modern smartphones will make such systems more affordable, accessible, plus easier to setup. Although current phones are not designed to determine their position other than with GPS, some of them have hardware components on board that could be used for this purpose as well. For example, every single mobile phone has speakers which can be used to generate sound as well as ultrasound. Sound travels relatively slowly through air, and by using the difference between when the time signal was generated at the phone and the time it was received at several locations it is possible to determine the phone's position through trilateration with high accuracy. Under ideal circumstances (e.g. perfect synchronisation, no interference, accurate calibration) this method can in theory deliver sub-centimetre accuracy. This appears to be very promising in contrast to most other available methods, in particular those that use electromagnetic signals (Wi-Fi, Bluetooth, GSM etc.), which can hardly reach meter accuracy due to hardware constraints. However, there are various sources of error that can degrade the accuracy of ultrasound positioning and it was therefore necessary to research how to mitigate or even eliminate them. For example synchronisation can be completely avoided by using a Time Difference of Arrival approach and precision can be improved by increasing the sampling rate.

There are a number of reasons why a software solution as the implementation method for the mobile positioning system was chosen as opposed to building custom hardware, such as RFID badges:

- Easier to purchase ubiquitous hardware. Users may use their own phones.
- No hardware engineering experience required to develop, install and maintain.
- Smartphones have touchscreen displays, which enable the user to interact with the system directly and as such benefit from any functionality the positioning system allows.
- Apart from locationing, specialised hardware such as RFID/Wi-Fi tags or badges cannot do much else, so they would have to be bundled or plugged into some other device to be useful. However, a mobile phone on its own is a full-fledged computing platform capable of hosting the client side of a complete LBS.
- Most smartphones are equipped with a magnetometer and accelerometers. Some even have a gyroscope. With some effort this hardware can be used to accurately determine the orientation of the phone in 3 dimensions. Combined with an accurate position it may be possible to determine what the phone is pointing at precisely. This will enable directional querying, which is one of the most useful and intuitive applications of a Location Based Service.

The above points demonstrate how the proposed solution can eliminate a gap that currently exists between accurate indoor positioning systems and services that could greatly benefit from such positioning. Because ultrasound indoor positioning has not been attempted yet on COTS mobile phones, sub-meter accuracy has so far only emerged in the form of tags and badges. Tags and badges have little or no feedback by definition. There is no visual or audio interface either. As such, while the positioning system may be aware of the user's location, the user himself has very little immediate benefit in terms of what an LBS could offer. At the same time there exists an area of context-aware services that could be greatly improved thanks to the combination of mobile phone hardware with accurate position and orientation. Some of these systems

already exist in museums such as virtual tour guides (Chou 2004; Tsai 2010), but remain very limited because specialised devices have to be kept cheap to replace in case they get lost or damaged. Other services have not even received a lot of exposure yet in interactive form such as indoor navigation systems or building evacuation procedures. The proposed method attempts to unite mobile phone interfaces and accurate indoor positioning so that the development of services such as those mentioned above is not limited by hardware cost, but is only a matter of writing software applications.

As an example following use case can be considered. The user enters a history museum. There is a notice at the entrance that says the museum is equipped with a smartphone virtual tour/navigation infrastructure. There is a QR code under the message. The user scans the QR code and installs the required app. The system enhances the user's experience in the following cases:

1. The user notices a small 12<sup>th</sup> century vase and decides to learn more about it. He points the phone at the vase and initiates directional querying. The vase is located on a shelf in a row among other vases. Thanks to high accuracy of positioning the system is able to identify the correct vase. The smartphone displays a page with a brief description of the exhibit followed by links to relevant text, audio, video and interactive material. Thanks to flexibility of the app the user is able to choose content that he is interested in. The user looks at pictures from the excavation site and even watches a video of the exhibit being excavated taking advantage of the powerful multimedia capabilities of the device he owns.
2. The user notices a painting he has seen before. He does directional querying and finds out it is by Claude Monet an impressionist. The user decides to read more about impressionism and accesses a Wikipedia page through a 3G connection on his mobile phone. He decides that he wants to see more paintings by impressionists

and requests the museum app to guide him. The system generates the most efficient route such as to cover all the required paintings starting at the user's current location. Thanks to high positioning accuracy the system is able to show on the screen the relative position the next painting to the user, making navigation more comfortable and avoiding confusion. Unfortunately some notable impressionists are not represented in this museum, however the user is able to become familiar with their works through high-resolution digital versions of their paintings being displayed on the screen of his smartphone.

3. The user decides he needs to find a bathroom. He presses a corresponding button in the app. The system finds the closest bathroom and guides him there by displaying a map, current location, route and directions.
4. The user realises that he immediately needs assistance from museum staff. He presses a corresponding button and is informed that he will be assisted shortly. The system finds the nearest staff member who also has a smartphone with a staff version of the app. He is quickly guided to the location of the user.

Altogether the new museum experience combines immersion and tangibility of a traditional museum visit plus the flexibility, richness and comprehensiveness of online surfing. While other features of the system make the visit more comfortable and safe.

### 1.3. Aims of the Thesis

Sending and receiving ultrasound signals is an unconventional use of sound hardware both in mobile phones and computers. It is therefore necessary to find out how well mobile phones can produce ultrasound, whether some audible noise will be produced alongside ultrasound and if necessary what can be done to avoid that. It will be necessary to find microphones that can detect the mobile ultrasound signal, but further

research into microphone limitations is not necessary because, unlike phones, microphones are not introduced into the system by the user. Also the range at which the signal can be reliably detected will have to be identified in order to determine coverage area.

Synchronisation was identified to be a major source of error for a locationing system that relies on time-of-flight (Bahl 2000; Peng 2007). We are therefore motivated to develop an asynchronous trilateration procedure that uses Time Difference of Arrival (TDOA) techniques.

Unless further obstacles are identified in the process, the above actions should be enough to gather the necessary information needed to build a working prototype of an accurate ultrasound indoor positioning system for mobile phones. This system will be tested to determine accuracy, reliability and possible shortcomings.

The above objectives can be summarised in the following research questions (RQ):

RQ 1: Can ultrasound be reliably reproduced by mobile devices?

RQ 2: What are the desirable characteristics of the emitted signal?

RQ 3: What is the maximum distance at which an ultrasound signal emitted by a mobile phone can be reliably detected with a microphone?

RQ 4: Can ultrasound positioning be done asynchronously?

RQ 5: What accuracy can mobile asynchronous ultrasound trilateration offer?

RQ 6: What impact does orientation of the speaker and the way a user stands have on accuracy and reliability?

RQ 7: Can background noise cause false positives and how can this be countered?

Answers to these questions will produce the following original contributions to the field of indoor positioning:

- **Ultrasound indoor positioning for mobile phones** – a positioning approach with sub metre accuracy that works with unmodified off-the-shelf mobile phones, doesn't rely on presence of rare, experimental or emerging hardware, and is very lightweight on the phone side.
- **Asynchronous trilateration** – a method derived from traditional least squares trilateration that eliminates the need for synchronisation at the expense of increasing the minimum number of required control points (microphones) by one. As it relies on least squares to calculate the most likely position, extra microphones can be used robustly to improve accuracy and reliability.

## 1.4. Document Structure

The rest of the chapters are organised as follows:

Chapter 2 covers literature review and is divided into “Positioning Methods”, “Services” and “Interfaces” followed by a summary. Positioning methods represents the largest section and in turn is divided into a number of subsections each covering a particular method. These are followed by a separate summary.

Chapter 3 covers preliminary experiments. Sections 3.1 and 3.3 cover experiments with ultrasound, the ability of unspecialised sound hardware to produce and receive ultrasound at various distances, angles and frequencies. Section 3.2 covers signal design.

Chapter 4 covers development and testing of a novel Time Difference of Arrival (TDOA) algorithm. Section 4.1 introduces the concept of using Least Squares for Time of Arrival (TOA) trilateration. Section 4.2 introduces our new positioning algorithm and explains how it is derived. In Section 4.3 the new algorithm is tested on paper.

Chapter 5 covers evaluation of our positioning system prototype in a real-world environment. Sections 5.1 – 5.3 cover a range of experiments designed to test accuracy and reliability of the system. Section 5.4 discusses impact of background noise on our mobile positioning prototype.

Chapter 6 starts with a summary of work followed by thesis contributions. Future work is the longest section and is divided into four subsections. Each subsection is dedicated to a particular research direction. The final section contains the overall conclusions.

## 2. RELATED RESEARCH AND BACKGROUND

This chapter reviews the current literature and gives an overview of the related work in the field of *indoor positioning* as well as *Location Based Services (LBS)*. LBS applications are interesting to the work in this thesis because it is an area where indoor positioning can be applied. Although positioning is a prerequisite for any LBS, the two fields need to be researched separately because currently the field of mobile indoor LBS is not very well developed. Whether this is because of the lack of a universal platform or the lack of functionality is an open question. A lot of indoor positioning solutions have never been used in conjunction with LBS usually because of the platform constraints (e.g. lack of screen, keyboard, operating system) and because they were developed to track people or objects rather than to provide some kind of context-aware information service. It is therefore necessary to review the underlying technologies of such systems in order to establish whether they can be used on a platform with fewer constraints and even form the universal platform mentioned above. The trends in Location Based Services need to be researched in order to form requirements for a system that will supply positional data. Because the interfaces used by different LBS are of particular importance for proper data display and user-friendly interaction, they are also covered in this chapter.

The chapter is divided into three sections: Positioning Methods, Services and Interfaces. Each section is concerned with a particular question. In case of Positioning Methods the question is “How is position information gathered?”. For the Services section the question is “What service is delivered?”, while for Interfaces “How is the service delivered?”. The combination of all three aspects defines how useful and valuable an LBS is. After all, practically any LBS can be replaced with a non-context-sensitive tool. A navigation program can be replaced with a map. A virtual museum guide can be

replaced with a brochure. It is how more convenient and informative a user's experience becomes thanks to an LBS that is important.

## 2.1. Positioning Methods

This section is dedicated to positioning methods that are applicable to today's off-the-shelf smartphone hardware. Each subsection is dedicated to a particular kind of hardware that is found in all modern smartphones, positioning methods that rely on this hardware as well as limitations of the underlying technology are discussed and compared.

In terms of the impact on user experience the most important parameter of a positioning method is accuracy. Very poor accuracy may restrict the choice of applicable interface types, while better accuracy can improve service usability, provided an interface allows for that. For example, an interface that determines which wall the phone is pointing at and returns a list of all objects on that wall does not need centimetre accuracy, but for retrieving a specific book off the shelf in a bookstore or library does.

### 2.1.1. Satellite Navigation Systems

**Global Navigation Satellite Systems (GNSS)** use radio frequency (RF) signals regularly sent by a constellation of satellites to estimate position of the receiver. Up to ten years ago, it was the only practical system which could successfully use the time-of-flight property of an RF signal. Since it travels at the speed of light, just like other subsets of electromagnetic radiation, it requires extremely precise measurements. Today this is not the case, as we will see later.

## 2 RELATED RESEARCH AND BACKGROUND

In the case of Navstar GPS, currently the only fully operational GNSS, satellites have an average orbit altitude of 20,200 kilometres above the surface of the earth. The system consists of three segments. The Space segment consists of 24 Navstar satellites, and the Control segment encompasses a number of facilities located in different geographic locations with the Master Control Station (MCS) in Colorado Springs. The MCS functions include control of satellite station-keeping manoeuvres, reconfiguration of redundant satellite equipment, regularly updating the navigation messages transmitted by the satellites, and various other satellite health monitoring and maintenance activities. Finally the User segment consists of receivers specifically designed to receive, decode, and process the GPS satellite signals. Each satellite broadcasts a 30 second long message at 50 bits per second. Each message contains the precise time it was sent, precise orbital position (the ephemeris) and general information about all the satellites and their orbits (the almanac). The almanac helps the receiver determine which satellites to listen to. Since it does not fit into one message it is split into several parts, which are continuously broadcast one after another in a loop. Previously this was a major reason for the startup delay and was addressed by Assisted GPS (AGPS). Modern receivers are much better at searching for satellites, so not having an almanac is no longer an issue.

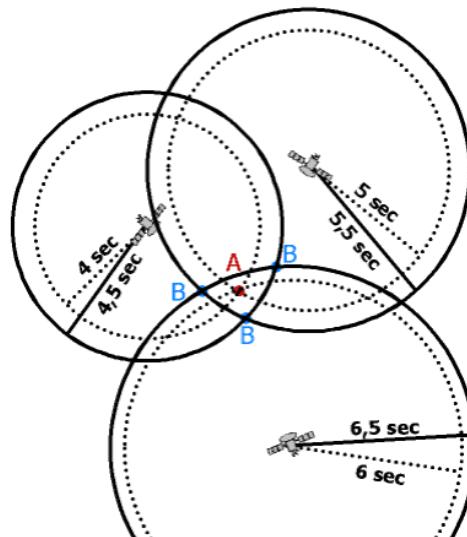
A GPS receiver determines the travel time of a signal from a satellite by comparing the "pseudo random code" the receiver is generating, with an identical code in the signal from the satellite. If we have this information for signals received from 4 different satellites, we know that the receiver is somewhere near the intersection of four spheres

## 2 RELATED RESEARCH AND BACKGROUND

with radius equivalent to the distance each signal travelled<sup>5</sup>. Since the ephemeris was encoded inside each message, it is possible to calculate the receiver position through a trilateration procedure. Theoretically, if a receiver clock was perfectly accurate 3 satellites would be enough to determine position in 3D space, as the three spheres would intersect in exactly two points, one on the earth surface precisely where the receiver is and one in outer space . Unfortunately, quartz clocks in GPS receivers are far less accurate than atomic clocks in GPS satellites and need to be corrected every second, so there are four parameters that must be estimated: the 3D coordinates of the receiver and the receiver clock error (Bossler 2002). If only three satellites are used the position will be calculated incorrectly and this error will go undetected. However if a fourth satellite is introduced, it becomes possible to detect and correct receiver clock error. Position determination in 2D with clock error is illustrated on Figure 1. The dotted lines show distance before clock correction (pseudorange) and the solid lines show distance after correction. A solution for three-dimensional space would require four satellites instead of three.

---

<sup>5</sup> "Navstar GPS User Equipment Introduction" Retrieved 12 November, 2012, from <http://www.navcen.uscg.gov/pubs/gps/gpsuser.pdf>.



**Figure 1: 2D position determination with 3 satellites and corrected clock error, adapted from GPS**

**Explained: Position determination<sup>6</sup>.** Solid circles represent distance to satellites according to the receiver. Because the receiver clock is inaccurate the three circles do not intersect at one point (B). After the clock is adjusted so that all three circles intersect at one point (A), the estimated distances are known to be more accurate. These are represented by dotted circles.

The major sources of range error for GPS are:

1. **Ionospheric delay.** Atmospheric factors give the biggest error and can be as high as 60 meters. Because the effect varies with signal frequency, it is possible to measure it by comparing L1 and L2 GPS signals (one is for public and one encrypted for military use). Also the same condition usually spreads over a large area and does not change very fast so it is relatively easy to keep track of it and broadcast the correction values to receivers.
2. **Tropospheric Delay.** Humidity also causes a variable delay which is more localized and changes more quickly than ionospheric effects, and is not frequency

---

<sup>6</sup> "GPS Explained: Position determination." Retrieved 12 November, 2012, from <http://www.kowoma.de/en/gps/positioning.htm>.

## 2 RELATED RESEARCH AND BACKGROUND

dependent. These traits make precise measurement and compensation of humidity errors more difficult than ionospheric effects.

3. **Ephemeris Error.** This error is the difference between the actual satellite location and the position predicted by satellite orbital data. Normally, errors will be less than 8 metres.

4. **Satellite Clock Error.** This error is the difference between actual satellite GPS time and that predicted by satellite data. This error is normally less than 6.5 metres.

5. **Multipath effect.** Multipath effect occurs when a signal bounces off the ground/buildings and arrives a bit later than the original signal. This problem is more severe in urban areas (2-4 meters error).

There are a few smaller error sources which are connected with hardware limitations and physical properties of the signal<sup>7</sup>. Most of them are discussed later in CDGPS subsection. Experiments have shown that for unassisted GPS the average position error ranges from 2 meters in an open area to 15 meters even in wide streets with four story buildings on both sides (Modsching 2006). In the end there is a need for a direct line-of-sight between the receiver and the satellites, which is the most serious limitation of using GPS in terms of this research.

There are a number of technologies that seek to improve core GPS performance. **Assisted GPS (AGPS)** improves startup performance by providing additional data, such as almanac, over an internet connection. Most commonly AGPS is found in GPS-enabled smartphones. Because in urban areas the signal will often bounce off buildings

---

<sup>7</sup> "Navstar GPS User Equipment Introduction" Retrieved 12 November, 2012, from <http://www.navcen.uscg.gov/pubs/gps/gpsuser.pdf>.

## 2 RELATED RESEARCH AND BACKGROUND

or fade while passing through tree cover, Time-to-First-Fix (TTFF) can be longer, so AGPS may be a useful improvement. Some users nevertheless dislike AGPS and even find it inferior to unassisted GPS because it uses their mobile internet traffic and therefore incurs costs. Also most devices fail to fall back to normal GPS mode which makes it impossible to determine location in areas with poor network coverage<sup>8</sup>.

**Differential GPS (DGPS)** is an enhancement to Global Positioning System that uses fixed ground-based stations to improve accuracy. These stations continuously monitor GPS signals and compare the results with their known real position, which was calculated with great precision. The resulting offset is broadcast via an Ultra high frequency (UHF) modem. In the radius of a few hundred kilometres all GPS receivers are going to have almost identical positioning errors, so when this offset is received and applied to roving GPS receivers, the resulting accuracy will be significantly better. DGPS is able to correct all errors mentioned in the above list except multipath. This means that in a radius of 100 km of the base station, in open areas error will be less than a meter and in an urban area no more than 6 meters. It has been reported that the error grows at a rate of 0.22 meters per 100 kilometres (Cobb 1997) (Badea 2005). Unfortunately, at present DGPS is not available on mobile phones.

**Carrier-Phase Differential GPS (CDGPS)** is able to deliver accuracy within a few centimetres. Normally GPS relies on matching pseudo random code to measure how long it took the signal to reach the receiver. Both receiver and satellite generate the same code simultaneously. When the signal is received, the receiver slides one of them

---

<sup>8</sup> "Navstar GPS User Equipment Introduction." Retrieved 12 November, 2012, from <http://www.navcen.uscg.gov/pubs/gps/gpsuser.pdf>.

## 2 RELATED RESEARCH AND BACKGROUND

until they sync up. The amount it has to slide the code is equivalent to the signal's travel time. The problem here is that the pseudo random code has a bit rate of about 1 MHz. Even if the signals are perfectly phased, this still allows for an error of a few meters. CDGPS aims to reduce that error by utilising the frequency of the carrier signal, which has a cycle rate of over 1 GHz. A receiver can measure the carrier cycle up to a fraction of a percent, but the number of whole cycles has to be derived indirectly (O'Connor 1997). This is known as integer ambiguity. Until recently, determining these integers has been a cumbersome and time-consuming process. It was either necessary to start at a precise location or wait for an extensive period of time. Some solutions required calculating a trajectory which is suitable for planes and cars, but is not practical for pedestrian and indoor use (Cobb 1997). Thanks to advances in computational power of mobile devices and CDGPS technology this is no longer a major problem.

**Pseudolites** are transceivers that send the same kind of signal as GPS satellites but are located on the ground. If positioned at fixed locations, ephemeris error and atmospheric delay are eliminated. If there are enough pseudolites at one location, usually four, they can serve as a standalone local navigation system and can even be placed indoors. Together with the centimetre precision of Carrier-phase positioning this is a very promising approach for indoor use. Unfortunately, it presently has a number of disadvantages. The first one is the price and availability of pseudolite units. In 2005 they cost 1,000-1,500 euro each and were discontinued shortly thereafter. This is a major reason why not much practical research involving full indoor “constellation” setups has been done to date, and why there is hardly any data on how well pseudolites work when there are obstacles, such as walls, in the way of the signal. Another problem is that unlike satellites which use very expensive atomic clocks, pseudolites use TCXO that are much less accurate. Altogether the technology is very promising and in the near future may become a popular and practical indoor solution for specific environments,

but ubiquitous use seems unlikely, especially considering pseudolites tend to interfere with commercial GPS receivers that are not designed to support pseudolite signals (Borio 2011). This technology as well as CDGPS and DGPS, are unsuitable for the work of this thesis mainly because they are not available on GPS-enabled mobile phones (Cobb 1997) (Badea 2005).

### **2.1.2. GSM**

Global System for Mobile communication (GSM) is the most popular standard for mobile phones in the world. GSM is a cellular network, which means mobile phones connect to it by searching for cells in the nearest vicinity. A GSM base station is typically equipped with a number of directional antennas with limited range that define these cells. In Europe 900 MHz and 1800 Mhz frequency bands are used. These frequencies are licensed plus each is divided into a number of physical channels, which are distributed among cells in such a way that cells assigned the same channel are located far enough from each other as not to cause interference. The channel to cell allocation is a complicated process and takes a lot of careful planning, which is why existing cellular network structure does not change very often. Because channels have to be reused, just a channel frequency cannot be used to identify a particular cell. Cells can be of different size, usually depending on how populated an area is. As a result urban areas have better cell granularity and therefore allow more accurate positioning, compared to rural areas, in contrast to GPS, which is opposite. Some shopping malls even have their own base stations (Kolodziej 2006) (Otsason 2007).

In terms of hardware GSM positioning is implementable on practically any device so is an ideal solution for mobile phones. Plus the GSM module is always on unlike other radio interfaces such as Wi-Fi, Bluetooth, and GPS, which add extra battery consumption and take time to start up. Unfortunately in terms of positioning accuracy

## 2 RELATED RESEARCH AND BACKGROUND

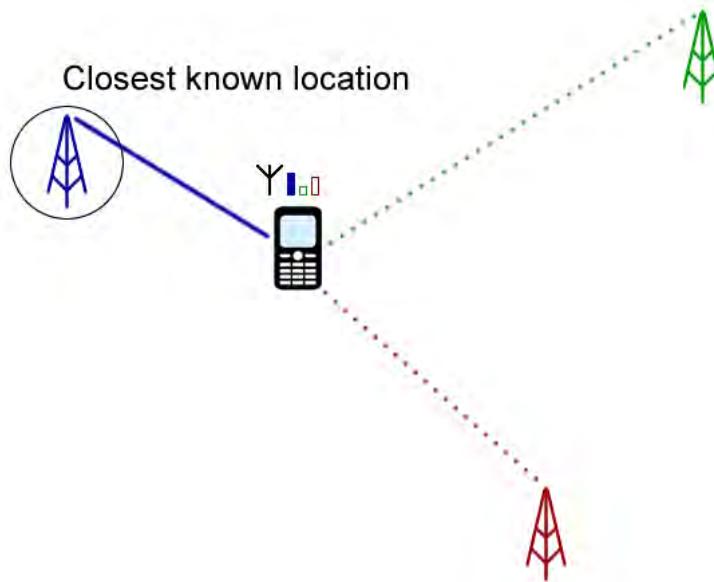
and complexity the interface is far from ideal. Unlike GPS, GSM was never designed for global positioning. What it offers is existing infrastructure and hardware that we can reuse and re-purpose (Otsason 2007).

GPS relies on very accurate clocks and other dedicated hardware specifically designed for the purpose of calculating the time-of-flight of a radio signal. Without such specialised hardware, using time-of-flight with RF signals is very difficult (Hoene 2008). Therefore other signal properties have to be used instead. Received-signal-strength (RSS) and bit-error-rate are most common. (We do not include Angle of Arrival as it requires specialised antennas.(Maddio 2010)) Theoretically there exists an inverse proportional relationship between the received signal quality and the distance it travelled (Thapa 2003). As the signal travels further it becomes weaker and will have a larger number of errors on arrival. An analogy would be taking a box full of Christmas balls and throwing it down a hill slope. At the foot of the hill we analyze the state of the box (RSS) and the number of broken balls (bit-error-rate) in order to estimate how long the slope is. In the case of GSM, RSS is one of the fundamental functions as the system needs to correctly track signal strength in order to know when to promptly switch between base stations. This at least ensures mobile phone hardware is capable of measuring RSS with fine precision. Although GSM adjusts the strength of transmission both at the base station and the mobile device, the broadcast control channel (BCCH), used to broadcast IDs of adjacent cells among other things, is transmitted at a constant power (Otsason 2007). However, depending on the operating system, access to some or most of this information is restricted for application developers. Because GSM positioning uses existing infrastructure and service, it is not viable to use bit-error-rate. Any data that we send via GSM will be charged by the service provider. Also we do not have control over the base stations and there are no specialised commands available to

developers, which would request a base station to echo back an unaltered message. Without such functions a scenario where an original message can be compared to the received message is hard to implement.

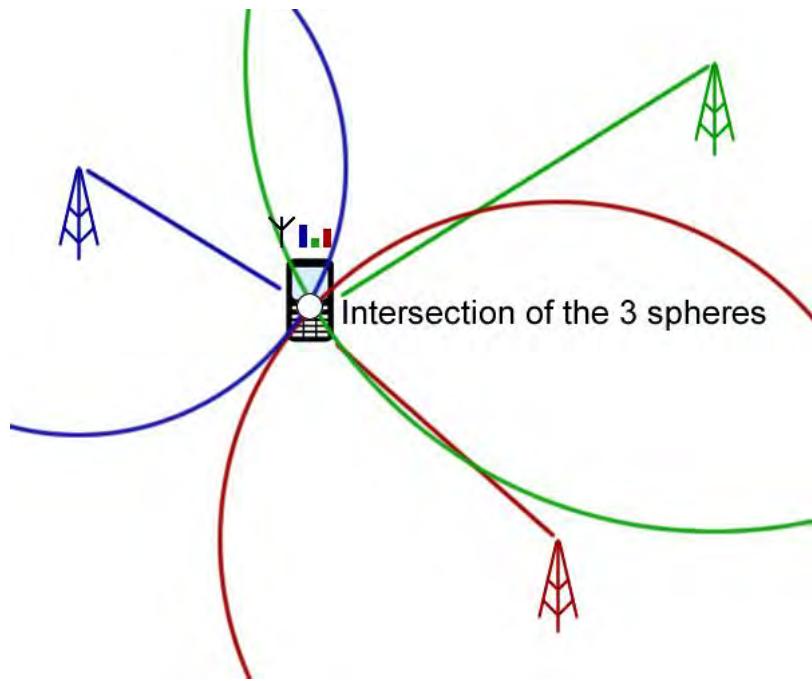
Generally, either one of the following techniques can be used to determine mobile phone location (Thapa 2003).

- **Nearest access point** simply tells the user, what is the location of an access point with the strongest signal, which most probably will also be the closest. RSS may also be used to determine how far away this access point is. An important advantage of this method is that it requires minimal computational power and is simple to implement. See Figure 2.



**Figure 2: Nearest Access Point.** The user's location is set to the location of the closest base station based on the received signal strength. The three bars above the phone indicate signal strength from the three strongest visible towers. The red and green bars are empty because they are visible but not connected.

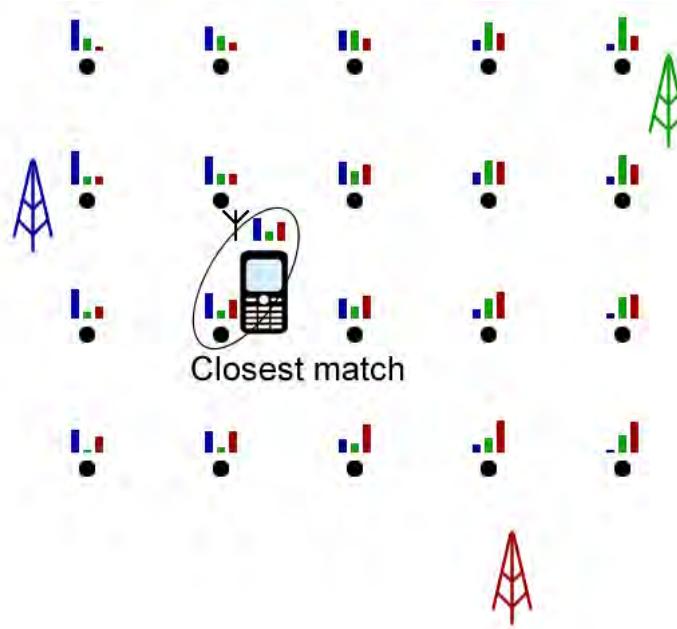
- **Trilateration** is a method of calculating intersection of three or more spherical surfaces given the position of their centres and radii lengths. Given the distance to three centres (signal sources) and their position it is possible to determine the receiver's position. Because signal properties do not very well correlate with distance, estimated position can have a significant error. Some technologies, such as Ultra wide band, are more suitable for this method than others. Using access points that are in very different directions is also an important condition. This method is often confused with the term triangulation, which uses angles of arrival. See Figure 3.



**Figure 3: Trilateration.** The user's location is assigned the location of the intersection of three overlapping spheres resulting from trilateration of the signal strength of the three strongest towers.

- **RF Fingerprinting** relies on recording the strength of several signals in multiple locations. The more signal sources are available, the easier it is to distinguish between two close calibration points. How far apart the calibration points are is also important. Generally, more readings give better precision, but at some point

better granularity stops affecting accuracy as the difference between nearby points becomes too subtle. During the use phase, current readings can be either directly compared with each fingerprint (raw measurements) to determine one's location, or with an observation model generated with the help of Gaussian processes (Ferris 2006). See Figure 4.



**Figure 4: RF Fingerprinting.** The black dots represent readings of received signal strength from the three strongest towers recorded in advance in a number of known locations. The received signal strength at the current user's location is matched against the recorded signal strengths and the location of the best match is assigned to the user's location.

GSM is one of the easiest technologies to implement **Nearest access point** with. As mobile phones automatically switch to base stations with the strongest signal, all that is left for the developer is to check which base station the phone is currently connected to and how strong the signal is. In operating systems with strong security and computational restrictions this may be the only possible method to determine location. The difficult part is to collect the existing locations of the cell phone towers.

## 2 RELATED RESEARCH AND BACKGROUND

Depending on the country and the particular service provider it may be very easy or near impossible to collect such data from published sources. If impossible, the only way left would be to collect this data manually which, considering the very poor accuracy of this method, would not be an effective solution.

Just like the previous method, **trilateration** requires known locations of base stations. The phone needs to have access to IDs and signal strengths of at least three nearby towers, which should not be a problem for most modern mobile phones. There are two software platforms that use cell tower trilateration to some extent. In Skyhook<sup>9</sup> service, cell tower trilateration is used as a coverage fallback, when neither GPS nor Wi-Fi positioning are available. Skyhook claims their GSM trilateration provides 200 - 1000 meter accuracy. Another platform is Navizon<sup>10</sup>, which combines Cell and Wi-Fi trilateration. Because base stations are fixed, there is nothing to improve on in order to make GSM trilateration more suitable for indoor use, and with such accuracy limitations it is unsuitable for this research.

On the other hand, **RF Fingerprinting** can be used both for outdoor and indoor positioning. In fact, it depends on how many fingerprints we are prepared to collect and maintain. Outdoor fingerprinting is an excellent alternative to trilateration. Ferris et. al. collected training data for three service providers over an area of 465 square kilometres with the help of a GPS unit, while driving a car. Results showed that GSM

---

<sup>9</sup> "Skyhook Wireless: How it works" Retrieved 12 November, 2012, from  
<http://www.skyhookwireless.com/howitworks/>

<sup>10</sup> "Navizon Technical Paper" Retrieved 24 May, 2009, from  
<http://www.navizon.com/navizon-how-it-works>

## 2 RELATED RESEARCH AND BACKGROUND

fingerprinting can give a median error from 293 meters in a suburban area to as low as 94 meters in the city centre. These results were compared to their Gaussian Processes (GP) (see pages 25-26) method which gave 236 and 128 meters respectively (Ferris 2006). Because pure fingerprinting is unable to localize in areas that were not covered during the training phase, it is much more suitable for car navigation. GP locationing however will deteriorate only slightly as the user wanders into an area some distance away from a major street.

Indoor fingerprinting is different because it appears possible to thoroughly collect RF fingerprints in the entire building with any granularity that appears practical. Otsason et. al. showed that indoor GSM fingerprinting can achieve median accuracy of 5 meters in a large building and even be able to differentiate between floors (Otsason 2007). Their method relies on wide signal strength fingerprinting. In addition to the 6 strongest cells, they recorded readings of up to 29 more cells that were strong enough to be detected, but too weak to be used for communication. Measurements were taken between 1 and 1.5 meters apart using a GSM modem. Their system proved to be effective in a number of environments such as a wooden house and a large concrete building, while being able to correctly identify floors 89% and 97% of the time. Median accuracy ranged from 2.48 to 5.44 meters. Among GSM positioning methods, wide signal strength fingerprinting seems the most promising approach for indoor use and combining it with Gaussian Processes, presented by Ferris et. al. (Ferris 2006), could prove an interesting approach. While it probably will not significantly increase accuracy, it should allow for reducing the number of fingerprints taken during the training phase without any significant drop in positioning accuracy. Unfortunately it is important to note that wide signal strength fingerprinting has only been implemented with a GSM modem, that can detect and use up to 35 cells, as unfortunately commercial phones limit access to

information from the 6 strongest cells in the best case or even only one cell in the worst case.

### 2.1.3. Wi-Fi (802.11)

Wi-Fi is a certification mark developed by the Wi-Fi Alliance<sup>11</sup> to indicate products that are based on the Institute of Electrical and Electronics Engineers' (IEEE) 802.11 standards (Kolodziej 2006). Most of the time Wi-Fi certified hardware, for example a laptop, is used to connect to an access point (Wi-Fi router), in order to get an internet connection. Client-to-client connection is also possible and is called an ad-hoc mode. A good example of successful use of this mode is peer-to-peer multiplayer on Nintendo DS<sup>12</sup>, a mobile gaming console. It is sometimes possible to make two Wi-Fi certified devices work in an ad-hoc mode even if this was not anticipated by the manufacturers, but it is very challenging to setup and is not possible on many devices. Currently the Wi-Fi Alliance is working on a new specification called Wi-Fi Direct<sup>13</sup>, which will overcome this problem. The new specification can be implemented in any Wi-Fi device from mobile phones, cameras and notebooks to keyboards and headphones. When operational, any one of these devices will be able to establish a client-to-client connection or connect in a group and advertise available services. Significantly, the new generation of Wi-Fi Direct devices will be able to create connections with millions of older devices already in use. This new technology is expected to have a huge impact on

---

<sup>11</sup> "Wi-Fi Alliance" Accessed 12 November, 2012, at <http://www.wi-fi.org>

<sup>12</sup> "Nintendo DS Official Site" Accessed 12 November, 2012, at <http://www.nintendo.com/ds>

<sup>13</sup> "Wi-Fi Direct Frequently Asked Questions" Retrieved 12 November, 2012, from [http://www.wi-fi.org/files/faq\\_20100916\\_Wi-Fi\\_Direct\\_FAQ.pdf](http://www.wi-fi.org/files/faq_20100916_Wi-Fi_Direct_FAQ.pdf)

## 2 RELATED RESEARCH AND BACKGROUND

the mobile device industry, make Wi-Fi even more pervasive than it is now and potentially replace Bluetooth in time.

In a way, Wi-Fi is unique among potential positioning platforms. On one hand Wi-Fi routers can be easily purchased and installed anywhere without restrictions or certification. On the other hand there is a massive existing infrastructure, that can be used both for indoor and outdoor positioning. In case a new indoor positioning system needs to be deployed, the environment may already have enough access points, for example an office building. In case of an outdoor system, most parts of a city will have numerous access points: free hotspots in city centre cafes, private access points owned by residents, etc. Unless a Wi-Fi access point was configured not to broadcast its SSID for security reasons, its signal strength and signature can be easily used as a reference. However, unlike GSM, this ease of installation leads to ease of mobility and thus you can never be sure that the recorded positions of the Wi-Fi routers did not change over time.

Commercial Wi-Fi positioning, though still being considered a novelty, has already proved to be a success. Navizon and Skyhook advertise their hybrid positioning services as an appropriate enhancement/replacement for GPS in urban areas. Both unite Wi-Fi positioning and Cell Tower triangulation plus optionally GPS to deliver fast and reliable positioning with 10-20 meter accuracy. Such high accuracy is entirely achieved thanks to Wi-Fi with the other two technologies serving as assistance or backup. The solutions are purely software based and require the phone only to be Wi-Fi compatible. Skyhook's system uses unique MAC addresses of access points and RSS of the signal broadcast by them. It is not clear from their documentation whether their positioning algorithm is based on trilateration or fingerprinting. In case of Navizon, their system uses trilateration. This procedure is executed on the phone so as not to compromise

## 2 RELATED RESEARCH AND BACKGROUND

client's privacy. As important as the positioning algorithms are, databases with the signal readings need to be constantly maintained, updated, and expanded. Skyhook uses vehicle-based signal scanning (wardriving) to collect raw signal data, while Navizon buys this information from its customers with GPS enabled phones.

In the previous section we made a distinction between systems that infer user's position from the position of base stations and distance to them and systems that use location-specific statistics. Unfortunately there is not much more that can be done with GSM for positioning applications. Compared to GSM, indoor Wi-Fi positioning allows much more flexibility, but if we want to approach the sub meter accuracy threshold, using just either of them will not be enough. First of all a training phase is necessary. While information collected during the training phase can be used directly, it has a number of problems. To begin with, due to various factors, it is very noisy, so it has to be filtered out. To do that properly is a challenge in itself. Secondly we need to predict/extrapolate the readings outside the observation points. Both of these problems can be addressed by creating a model. Generally there are two distinct approaches to creating models: **physical**, when we model signal propagation through the environment and **mathematical**, when the model is made with the help of mathematical modelling of distribution.

**Physical approach.** Access points can be located inside the premises and rearranged when necessary. Because both the user and the access points are inside the building, it becomes sensible to attempt predicting how exactly the signal propagates through the environment and make relevant adjustments to the trilateration process. This was addressed in detail in an established work about RADAR by Bahl et.al (Bahl 2000). Although their system used WaveLan, a pre-IEEE 802.11 technology, their findings,

## 2 RELATED RESEARCH AND BACKGROUND

that contributed to a median resolution of 2-3 meters, may also be selectively applied to modern Wi-Fi which operates at the same 2.4 GHz frequency:

- Signal strength more strongly correlates with distance than signal-to-noise ratio.
- Signal strength at a given location varies significantly depending on the user's orientation because his body may be blocking the signal's path. It is therefore necessary to record several readings per one physical location during the training phase. Bahl et. al. recorded signal strength facing 4 different directions, but there are examples of continuously taking readings while slowly spinning around in order to factor out the effect of orientation (Krumm 2004).
- Using the layout information of the building and the Cohen-Sutherland (Foley 1996) line-clipping algorithm, it is possible to build an accurate signal propagation model by computing the number of walls that obstruct the direct line between the access point and the user. Data collected during the training phase is then used to develop a model that accounts for both free-space loss and loss due to obstructions.
- After a certain threshold, increasing the number of readings done during the training phase will gain hardly any accuracy. In the given case it was observed that for a 43.5 by 22.5 meter site, reducing the number of training points from 70 to 40 did not affect accuracy in any significant way.
- Transmitted signal will usually reach the receiver via multiple paths (e.g. multipath phenomenon). The strongest variation of the signal may be the one that reached the receiver via the line-of-sight, but not necessarily, because the path with least resistance may not be necessarily the same as the line-of-sight.
- When the distance between an access point and a receiver is long, free-space path loss dominates the loss due to obstructions.

## 2 RELATED RESEARCH AND BACKGROUND

- Parameters that were used to model wall resistance in the signal propagation model were similar across different access points. This suggests that walls in the same building impose the same resistance.
- It may be necessary to carry out the training phase at different times of the day in order to account for the varying number of people in the building, as their bodies weaken the signal.

A much more recent study was done by Mestre et. al. where 3 meter accuracy was achieved and site survey could be avoided altogether (Mestre 2011).

**Mathematical approach.** A signal strength observation model can be created even without signal propagation modelling. This is possible with the help of mathematical modelling of RSS distribution. Ferris et. al. has shown that this can be accurately done with Gaussian Processes (GP) (Ferris 2006). GPs are non parametric models that estimate Gaussian distribution over functions based on training data. The likelihood of a signal strength is extracted from a GP that is learned from calibration data. There are a number of reasons why GPs are a good choice for this task: they do not need an environment model, and can estimate the possibility of not detecting an access point as well as provide uncertainty estimates. A significant finding was that GP is able to accurately extrapolate the model into rooms that were not covered during the training phase. The average error of the system is roughly 2 meters.

An important part of these indoor positioning systems is prediction of user movements. The most basic piece of information we could use is whether the user is moving at all. This can be done with accelerometers, which are available in most modern smartphones. In the case of RF tags, accelerometers could be installed in the tags just for this purpose. This information can be used to put other hardware in a sleep mode in order to save

## 2 RELATED RESEARCH AND BACKGROUND

power. Also, because we know for sure the user did not move since the last location change, it could be used to prevent false readings. Making predictions based on realistic human walking speed can be a great approach to correcting positioning errors. For example if a reading suggests that a person can be either 5 meters away or 20 meters away from the point known 5 seconds ago, the first case should get a higher probability. It is important to make this as flexible as possible, because any number of previous readings could be inaccurate. It is not uncommon to use this approach as a core of a positioning algorithm (Krumm 2004).

Research has been done on using Time of Arrival with Wi-Fi by Hoene et. al (Hoene 2008). Their approach is software-based and works with regular WLAN chipsets. The software is called “Goodtry”. It relies on access to timestamps generated either at the time of sending or receiving packets, which some chipsets do not provide. While the resulting four meter accuracy is very impressive, provided Wi-Fi hardware doesn’t support TOA measurements, at the moment it doesn’t have accuracy advantage over more conventional Wi-Fi based methods.

Currently one of the best commercially available solutions for indoor use, Ekahau<sup>14</sup>, can track RF tags with accuracy of 1-3 meters. This is at least twice as good as the best GSM indoor positioning known.

Wi-Fi positioning is a very good choice for indoor use. It has been thoroughly researched and even implemented by a few companies on a commercial basis. With

---

<sup>14</sup> "Ekahau Real Time Location System" Accessed 12 November, 2012, at <http://www.ekahau.com/products/real-time-location-system/overview.html>

Google gradually expanding navigation indoors for Google Maps, which come preinstalled on all Android devices, Wi-Fi fingerprinting is likely to become a de facto standard for mobile indoor positioning (Ball 2012). Unfortunately there is very little hope that this technology will achieve stable sub-meter accuracy with current technology. We see a lot of research potential in this field when Wi-Fi Direct becomes widely available. Unfortunately at the moment of writing Wi-Fi Direct enabled devices were not yet available for purchase.

### 2.1.4. Bluetooth

Bluetooth is an open wireless protocol for data exchange over a short distance. It operates in the same 2.4 GHz band as Wi-Fi, but uses a weaker signal and implements adaptive-frequency-hopping in order to avoid interference with other devices in the same band. Bluetooth devices are categorized into three classes according to power. Class 2 with approximate range of 10 meters is the most common. Class 1 has a range of 100 meters and Class 3 a range of 1 meter<sup>15</sup>.

Bluetooth was targeted to enable wireless communication between mobile devices, hence a weaker signal for lower power consumption. The three most common uses for Bluetooth in mobile phones are connecting headsets in order to make hands free calls, connecting two mobile phones to exchange data (e.g. ringtones) and finally synchronisation with a PC. Today, Bluetooth is practically the only choice when it comes to wireless mobile phone headsets. Direct data transfer between two phones is

---

<sup>15</sup> "A Look at the Basics of Bluetooth Wireless Technology" Accessed 12 November, 2012, at <http://www.bluetooth.com/Pages/Basics.aspx>

## 2 RELATED RESEARCH AND BACKGROUND

not much used mainly because of the slow data transfer rate (3 Mbit max) and an unsophisticated interface. Now that almost every phone has some sort of internet connection it is faster and easier to send a file by email than to establish a Bluetooth connection. Bluetooth synchronisation with a PC is useful if the phone does not have a mini USB connection or internet use is either expensive or takes time to establish. One of the main problems with Bluetooth is speed, which was addressed in the 3.0 specification that allows two devices that attempt to exchange a lot of data to make use of the faster 802.11 technology in ad-hoc mode (Meyer 2009). This is an interesting development and may pose some serious competition to Wi-Fi Direct, but does not change the fact that in our age of High Definition video, Bluetooth is as good as obsolete. Nevertheless, Bluetooth hardware has become very cheap and is now ubiquitous. Theoretically it is a good choice for mobile positioning due to the flexible and open-ended nature of the protocol. Also, since there is such a multitude of devices with Bluetooth support, it would be desirable to use all of them as location reference beacons (control points). While reusing existing infrastructure (e.g. Bluetooth keyboard, mouse, etc.) is questionable because they are likely to change location, buying a number of such devices and using them as beacons is relatively inexpensive (Cheung 2006). For example, the cheapest Bluetooth headset now costs less than 20 Euros, an important advantage over competing technologies.

Unfortunately, Bluetooth does not possess a function that allows us to measure signal strength straight away. There is a command available through Host Controller Interface (HCI) called *Read\_RSSI* that returns a relative measure of RSSI and optimal signal strength range (Golden Receive Power Range, GRPR). If the signal is inside the range, the value is zero, stronger = positive and weaker = negative. Because the range is fairly wide, this command is least informative when the signal is within the optimal range.

## 2 RELATED RESEARCH AND BACKGROUND

The problem with using this command is that it is used internally by the two connected Bluetooth receivers. Once the value is not zero the two devices that share the link will attempt to adjust the transmission power level accordingly. It comes as no surprise therefore that experimentally no relation between RSSI and distance could be found (Hallberg 2003).

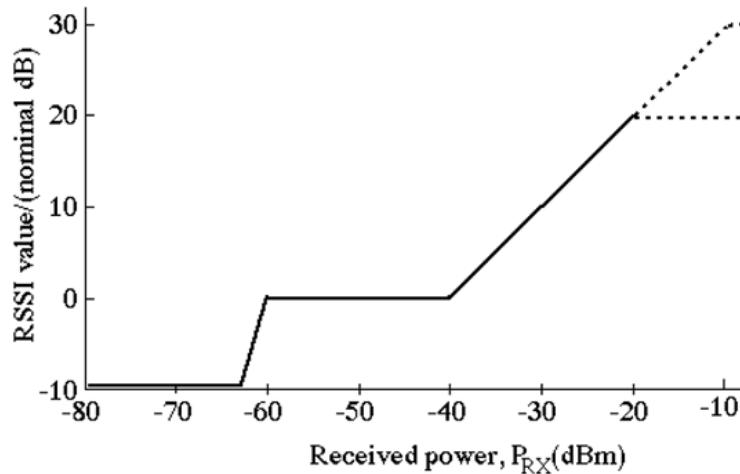
As discussed in (Hallberg 2003), transmission power level can be accessed with the *Read\_Transmit\_Power\_Level* command and in theory it should be possible to infer a phone's position with the help of these two commands. But it is not going to be very precise, because transmission power is not constant, so fingerprinting becomes a lot less reliable. For example, when we start to walk away from a beacon at some point in time *Read\_RSSI* will show that the signal is too weak and so transmission power may increase. Finally there is one more command *Get\_Link\_Quality*, which was shown to have somewhat more correlation with distance. Unfortunately this metric is manufacturer specific, which is why little research has been done with its use in positioning. Also it is important to note that implementation of commands such as *Read\_RSSI* is considered optional by many hardware manufacturers (Hallberg 2003).

It was shown experimentally by Zhou et. al. that it is possible to measure distance between two Bluetooth devices with a median accuracy of 1.2 meters by disabling the power control feedback system mentioned above (Zhou 2006). The received power was indirectly measured using the RSSI command, and the relationship between RSSI and actual received power for the devices used in the experiment is illustrated on Figure 5. From the diagram it can be inferred that it is best to adjust output power in such a way that RSSI values stay between 1 and 20. The two devices were programmed to take multiple readings of RSSI to check if its value was not zero, and change the power

## 2 RELATED RESEARCH AND BACKGROUND

output if it was. A Line-of-Sight propagation model can then be used to translate the collected data into distance. There is no information yet reported on how well this method performs with trilateration for determining position. A study done by Cheung et. al. suggests that when placed in a conventional office environment the reception field tends away from radial symmetry due to various factors such as electromagnetic noise, obstacles, etc. (Cheung 2006).

Currently it is impossible to disable the power control directly via the Host Controller Interface (HCI) on a mobile phone. Ability to force hardware to operate at maximum power has been introduced in Bluetooth 3.0. While this is better than an uncontrolled power level, the inability to manually adjust it is a major restriction. Theoretically it is still possible to replicate the above system for a mobile phone if the phone is only used to measure RSSI while all the necessary adjustments are done on the beacons. One problem we see with this approach generally is that Bluetooth devices that allow this sort of flexible programming do not have a self initializing ability present in headsets or computer mice. This can make setting up a large system very time consuming considering each beacon only has an effective range of up to 8 meters and needs to be initialised with each phone in question.



**Figure 5: Relationship between RSSI and Received power, adapted from Zhou S. (Zhou 2006).** The X axis represents the actual power of the signal when it is received. The Y axis represent RSSI value that corresponds to the given received signal strength.

Compared to other technologies, it is relatively simple to measure Bit Error Rate (BER) for a Bluetooth connection. There is an *echo* command that sends an arbitrary packet to a remote device and the remote device is expected to send back a packet with the same bit sequence (Thapa 2003). This command belongs to Logical Link Control and Adaptation Protocol (L2CAP), which is assumed to perform no error detection, so it should be possible to infer BER by comparing the initial message and the response. On many mobile platforms, applications are not expected to have direct access to L2CAP, and unlike HCI, it is manufacturer specific. Effectiveness of BER in Bluetooth positioning is also questionable because it has been observed that as the distance increases, the weakening of the signal manifests itself predominantly through failed synchronization rather than errors in the packet payload (Kolodziej 2006).

A very interesting feature of Bluetooth is the ability to form a network (piconet). A Bluetooth piconet can have 7 active slaves and up to 200 devices in parked mode. This is very useful for indoor navigation, because it means we could distribute a large number of beacons around the premises and a mobile phone could be connected to as

## 2 RELATED RESEARCH AND BACKGROUND

many as 7 beacons simultaneously (Kolodziej 2006). Because of the ambiguities associated with signal strength, the most efficient way to utilise a piconet would be a form of fingerprinting where we only record the probability of being able to connect to a beacon without going into details about the strength of the signal. If the beacons are arranged to overlap, combined with up to 7 simultaneous connections, it should be possible to at least achieve sub-room accuracy. While this solution is bound to have poor granularity, it is cheap to develop, is hardware independent, and can avoid many of the technical ambiguities and pitfalls connected with Bluetooth technology.

Although some studies suggest that Bluetooth positioning accuracy is on par with other technologies, there are a few issues that need to be taken into consideration (Hallberg 2003; Zhou 2006). The biggest problem is the time required to discover all possible devices in the current area. Due to the way Bluetooth works, it takes at least 10.24 seconds to “discover” in an error free environment when there are only 10 devices in the piconet (Kolodziej 2006). Considering it takes 16.7 seconds to cross a circle with a diameter of 20 meters at an average pedestrian speed there is a possibility that if we introduce more Bluetooth devices into the system, a beacon in a far corner that normally should be discovered will not have enough time to connect (Hallberg 2003). Some studies suggest that the lengthy discovery (inquiry) process can be avoided if the ID of target Bluetooth device is known and the two devices can go straight to the “paging” phase (Pals 2003; Subhan 2009). Another somewhat related problem is connected to security concerns. It is a widely accepted fact that it is dangerous to have Bluetooth enabled on a phone in a crowded area for an extended period of time. Apart from general wireless security threats such as eavesdropping, denial of services attacks, and man-in-the-middle attacks there is also a threat of Bluetooth related attacks, which can result in a hacker gaining access to personal information stored on the phone or even hijacking and making unauthorised telephone calls (Scarfone 2008). Many of the

## 2 RELATED RESEARCH AND BACKGROUND

serious Bluetooth vulnerabilities are a thing of the past now, but there is still some significant risk connected with this technology.

Bluetooth positioning is a good choice for coarse indoor positioning, e.g. room level. Finer accuracy is also possible but is connected with some technology-specific pitfalls. There appears to be no evidence that Bluetooth can reliably deliver sub-meter accuracy. We also find the idea of a headset piconet to be appealing as a cheap version of a positioning system infrastructure and are motivated to look into integrating it with a more reliable positioning technology, so that Bluetooth is used only to carry data as it was originally intended. Currently the most promising approach appears to be utilizing angle-of-arrival (Maddio 2010). Positioning with 20 cm accuracy was demonstrated by Nokia High Accuracy Indoor Positioning (HAIP) using Bluetooth 4.0 (Perez 2012). Unfortunately Bluetooth 4.0 will not be available on most mobile phones for a while.

### 2.1.5. Sound

Originally the primary purpose of a mobile phone was voice communication until smartphones turned into mobile computing platforms. By definition, a phone must have at least one microphone and one speaker. A combination of these two hardware components can be used to emit and receive sound waves in a similar way to how Wi-Fi/Bluetooth signals are sent and received with two major differences. One is that when working with sound on mobile phones, there are no existing notions of connections, packets, protocols or other features that come from networking to wireless communication. In other words we deal with “raw” sound. The other difference is the physical properties of sound.

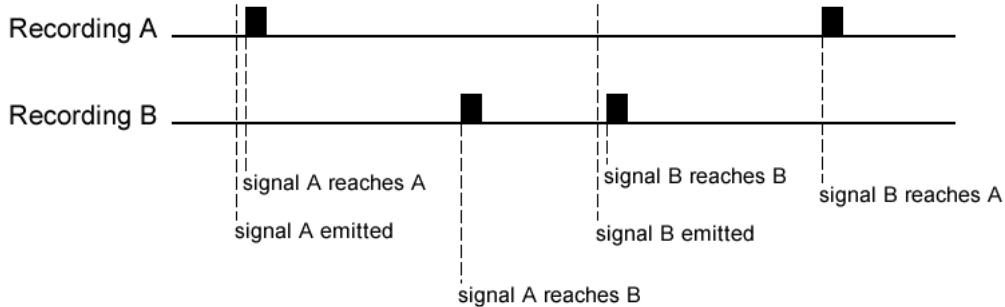
## 2 RELATED RESEARCH AND BACKGROUND

So far this chapter reviewed positioning methods that use electromagnetic waves. One of the main problems observed was that it is impractical to measure time-of-flight using unspecialised hardware because electromagnetic signals travel at the speed of light ( $3 \times 10^8$  m/s) and the distances we are trying to determine are relatively small. Sound, on the other hand, is a mechanical wave which travels at much lower speeds. In dry air at a temperature of 25°C the speed of sound is only 346 m/s. At such propagation speeds, one sample of a standard 44.1 kHz stream (44100 cycles/second) accounts for 0.8cm (Borriello 2005) (Peng 2007). In other words a signal will travel only 0.8 centimeters in the duration of the smallest time window available. Technically it is possible to work with sound even at 384 kHz, which can give much finer accuracy, however mobile phones normally don't support sampling rates above 48 kHz.

Unfortunately, an audio recording does not have a reference point for when the signal was sent, so a timestamp has to be collected therefore from the sender. If the sender and receiver have clock skew/drift between each other, this will result in synchronization uncertainty. One more uncertainty results from possible misalignment between the time a command to emit sound was issued and the actual emission time. Finally, receiving uncertainty occurs as a possible delay in the signal being recognised.

Peng et. al. showed that all of the above uncertainties can be eliminated when estimating distance between two devices (Peng 2007). Their “BeepBeep” ranging procedure involves two mobile devices starting to record sound before emitting short sound signals one after another. This way each recording has two reference points. Device A has a recording of the signal emitted by device A reaching the microphone on device A, and later of the signal emitted by device B reaching device A. Device B has a recording of the signal from device A reaching device B followed by the signal from device B reaching device B. The span between the two signals on device A is longer than on

device B since device A was the first one to emit sound. When the second span is subtracted from the first span the result is equal to twice the time it takes sound to travel between the two devices. (Figure 6)



**Figure 6: BeepBeep signal exchange.** The two horizontal lines represent recordings on each of the devices. Black boxes are actual sound signals that were recorded. The dashed lines represent events. Time interval between the two boxes on recording A minus time interval between the two boxes on recording B equals 2x the time it takes for the signal to travel between the two devices.

The “BeepBeep” procedure also involves adding the time that a signal travels between a speaker and a microphone on each device to the result. This is not strictly necessary, since the distance between a microphone and a speaker on each device cannot be longer than the length of the device itself and depending on the orientation of each in space may even introduce error. “BeepBeep” has presented itself very well in open environments, but unfortunately showed poor accuracy indoors at distances longer than 5 meters. Peng et. al. believe this was caused by the multipath effect(Peng 2007). The experiments were done in a small room with one or the other device close to a wall, which interprets a signal that bounced off a wall to be of comparable strength to one that arrived via the shortest path.

“BeepBeep” presents a very good idea that overcomes several problems common to acoustic ranging systems, but unfortunately the procedure is not very suitable for trilateration. To provide the necessary measurements, there have to be at least three or

## 2 RELATED RESEARCH AND BACKGROUND

four visible beacons which allows for measuring distance to them simultaneously either by listening to sound signals emitted by the mobile device or simultaneously emitting sound. It is argued that the first approach is better. Although it does not really eliminate any synchronization problems, many difficulties can be avoided by listening to just one signal at multiple locations. First of all, there is no need to distinguish between several different signals that arrive either simultaneously or very close to each other. Secondly, the computational load of trilateration will be on the server connected to the microphones, rather than the mobile device. As for using speakers as beacons, there are several ways to distinguish simultaneous signals. The simplest method is to use different frequencies. This is a good solution when not bound by a particular frequency range. Another solution is using coded pseudo-noise signals which, with properly chosen codes, can make signals orthogonal to one another and as a result easy to distinguish when overlapped (Peng 2007).

The effective range of transmitting beacons greatly depends on the volume of the signal and the direction of the speaker. As can be seen on Figure 7 for frequencies below 500 Hz sound propagation mostly follows a spherical model, however as frequencies become higher the spherical shape gradually starts to resemble a subcardioid (Figure 8, bottom), followed by cardioid (See Figure 8, left) and eventually supercardioid (Figure 8, right) closer to 16000 Hz (de Vries 1997). Cardioids among other shapes are traditionally used in acoustics to describe microphone response patterns, however specifically cardioids can be effectively used to describe sound propagation from a non-omnidirectional source at a variety of frequencies (Rumsey 2009). Data for sound directivity above 16 kHz is scarce and is not even supported by Common Loudspeaker

## 2 RELATED RESEARCH AND BACKGROUND

Format (CLF)<sup>16</sup>. Considering that at 40 kHz ultrasound generation also remains directional, the same can be expected at 20-22 kHz (de Vries 1997). Most smartphones have both a speaker and a microphone on the same side as the display screen while some also have a louder speaker on the opposite side. Regardless if the phone emits or listens for signals, beacons placed on the ceiling will have a direct line-of-sight with the phone's speaker/microphone while the user is using the phone. For small rooms it should be enough therefore to place a beacon at the top of every corner of the room. Unfortunately the cardioid shaped model suggests that if a room is significantly larger, the angle between a speaker and a microphone will be too great and the signal will fade too much, in which case a number of beacons will have to be placed on the ceiling to form a grid. Placing microphones flat against walls/ceiling should effectively counter the multipath effect, which speaks in favour of using the mobile phone speaker as the signal source.

It is evident from examples given above that the mobile device needs to communicate with the infrastructure from two points of view, first to communicate the intention to estimate position and secondly to exchange measurement results. It appears challenging to reliably transfer data with conventional speakers and microphones using a limited range of frequencies. According to research, the signal to noise ratio even at a range as short as 1 meter is too high to correctly decode more than 95% of the packets (Madhavapeddy 2003). Wi-Fi communication is a more reliable alternative. Therefore the sound signal can be of any length, shape and frequency as long as it can be reliably detected. Also there is no need to dynamically modulate the wave and it can be pre-generated and stored as an audio file. A signal length of 50 milliseconds was suggested

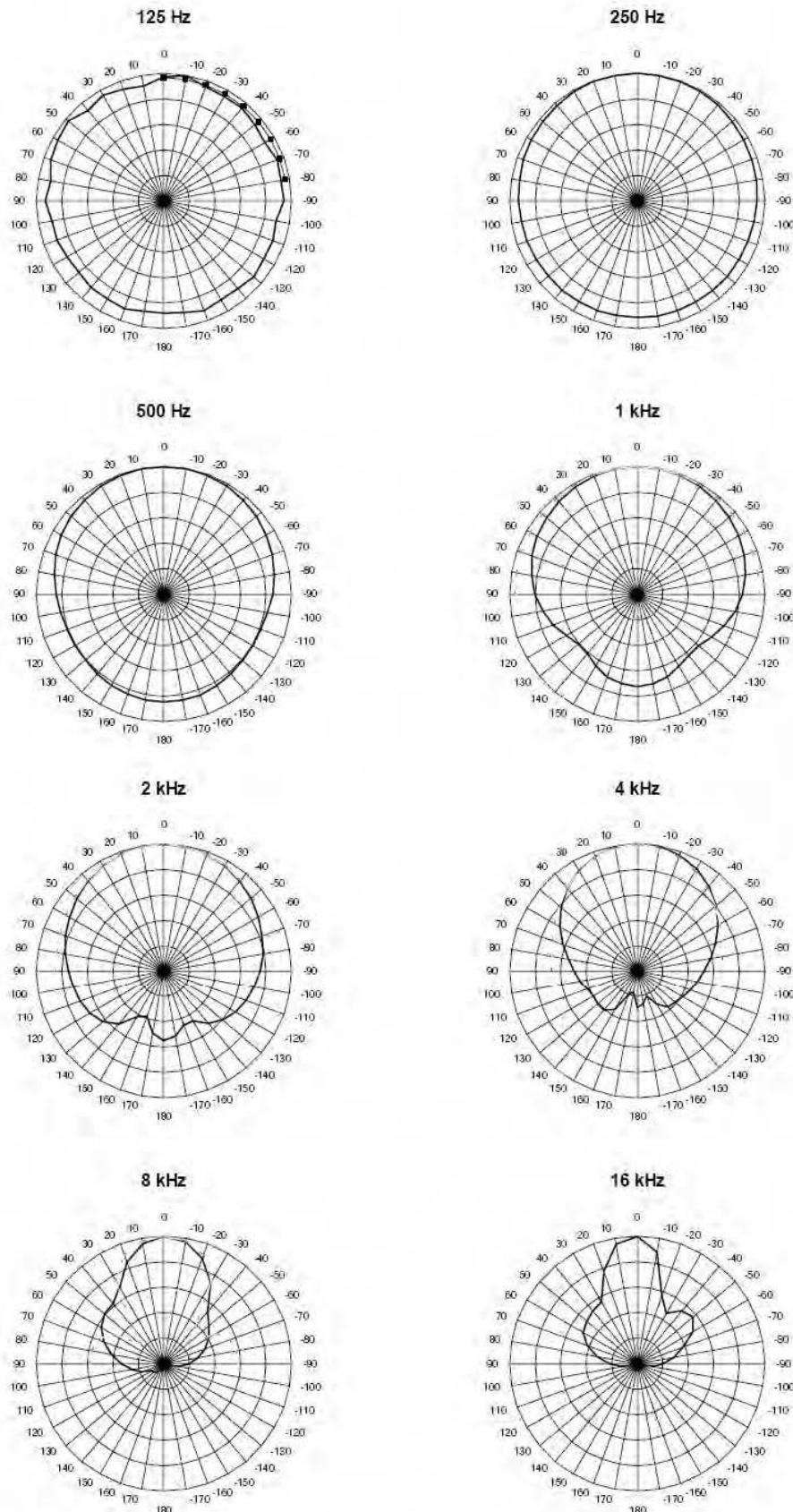
---

<sup>16</sup> "Common Loudspeaker Format" Accessed 12 November, 2012, at <http://www.clfgroup.org/>

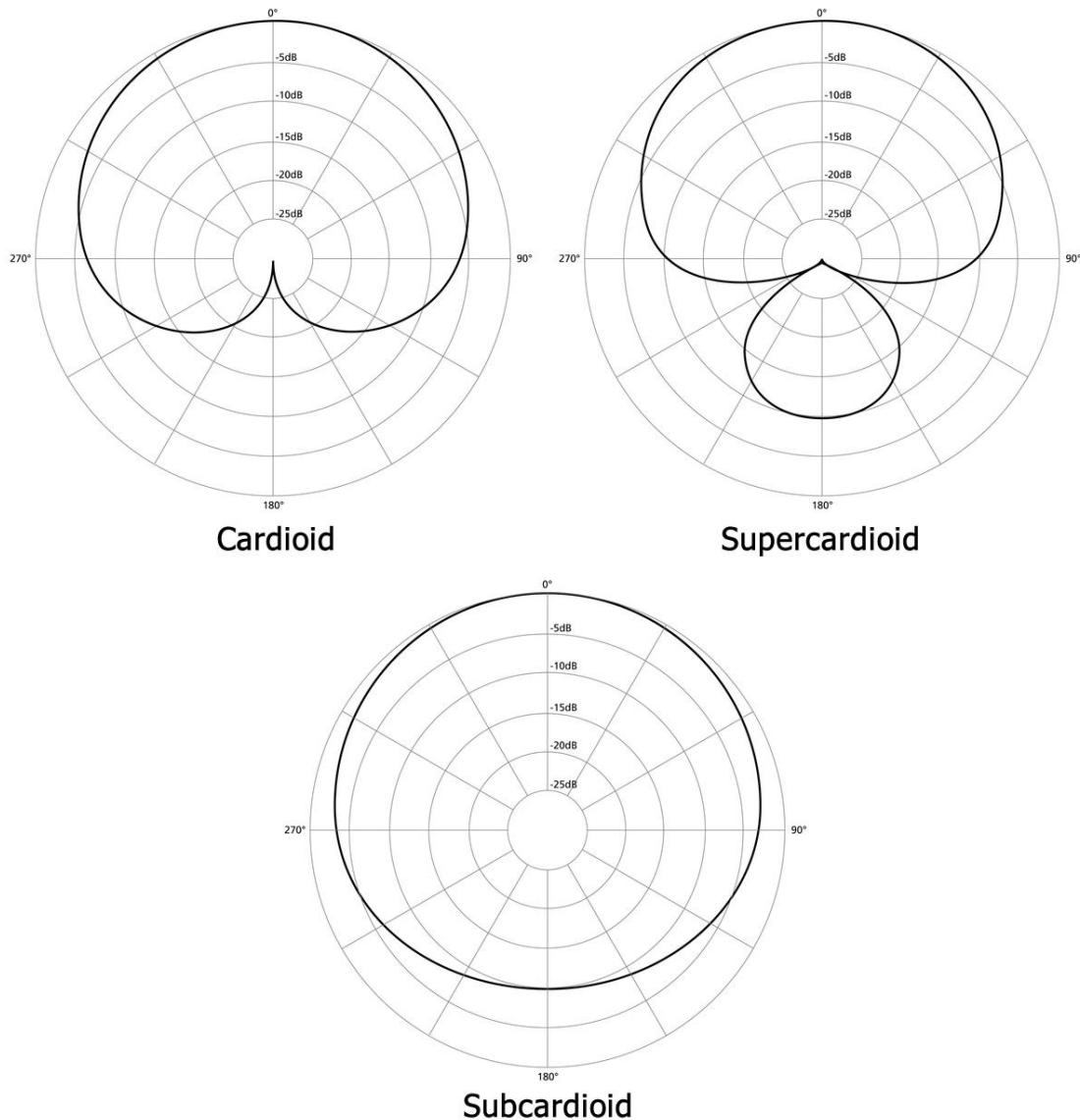
## 2 RELATED RESEARCH AND BACKGROUND

to be a good compromise between multipath effect suppression and noise resistance (Peng 2007). Although there are examples of successfully using even 10 millisecond chirps, one danger associated with using very short samples is that they can be easily masked by noise at large distances. It has been observed that the first few milliseconds of a sample playback are likely to come with a very large distortion which at certain frequencies appear to be a loud unpleasant click (Borriello 2005; Peng 2007). It is therefore recommended to gradually increase the amplitude of the signal. Regrettably, this may introduce some uncertainty to where the beginning of the signal is - an otherwise perfect candidate for a reference point. The end of the signal is a bad choice because it is likely to merge with an echo coming by an alternative path. The multipath effect is also the reason why it is not efficient to determine the middle of the signal and use that as a reference. One possible solution could be a signal that gradually increases in amplitude and immediately starts to decrease. This will form a “peak” that the receiver will try to detect. Finally the sound frequency presents a choice between efficiency and usability. It has been suggested that anything above 8 kHz attenuates too quickly (Peng 2007). On the other hand it appears desirable to use a frequency that is inaudible to humans. Frequencies above 20 kHz (ultrasound) generally cannot be picked up by human ear. While these frequencies reduce the effective range of our system, this is offset by a noiseless positioning system placing more importance on user experience. If necessary, this would justify an increase in the number of necessary beacons. Also higher frequencies are easily stopped by obstacles, while lower frequencies can even penetrate walls. If taken into account when designing the system either could be used to an advantage.

## 2 RELATED RESEARCH AND BACKGROUND



**Figure 7: A polar plot depicting 6dB/div measurement for a range of frequencies with a step of one octave, adapted from de Vries et. al. (de Vries 1997) *Direction in which the speaker is pointing is marked with zero (top of the polar plot).***



**Figure 8: Cardioid and Supercardioid polar patterns.** Adapted from Rumsey (Rumsey 2009).

There are many examples of indoor positioning systems that successfully utilise ultrasound. Holm et. al. (Holm 2009) use the fact that ultrasound cannot pass through walls in order to localise a user with room-level accuracy. Other approaches rely on trilateration to get very high accuracy.

- The Bat (Harter 1999; Addlesee 2001) relies on a grid of sensors attached to the ceiling. When a device, called a Bat, is triggered by an RF signal, it produces an ultrasound signal, which is then detected by several sensors. Time-of-flight is

## 2 RELATED RESEARCH AND BACKGROUND

calculated based on the time RF signal was sent and the time ultrasound signal was received at each sensor. Orientation as well as position of an object can be calculated by attaching two Bat modules. The Bat has accuracy of around 3 cm. In SNoW Bat accuracy was improved to 15mm (Baunach 2007). Also self-localization of nodes was introduced to aid faster deployment at the expense of accuracy.

- The Cricket (Priyantha 2005) also uses a grid of sensors and has similar accuracy. The main differences are that it is decentralised and that the mobile module produces RF “advertisements” simultaneously with ultrasound signals instead of being triggered by an RF signal. This way the system doesn’t explicitly track the user, which was done to improve user privacy. The Cricket has accuracy of around 3 cm.
- DOLPHIN (Minami 2004), similarly to SNoW Bat, is aimed for self-localisation. DOLPHIN nodes can act both as transmitters and receivers. This allows a strategy where only a number of nodes may know their location. Other nodes can discover or update their location through trilateration with master nodes. After a node’s position was determined it can participate in trilateration in the role of a master node. DOLPHIN has accuracy of around 15 cm.
- “High Performance Privacy Oriented location system” (Hazas 2006) uses wideband transmitters placed on the ceiling, which transmit their coded signals simultaneously at defined times. For each signal a unique gold code (Gold 1967) is used. The mobile device is aware of which gold codes are used and is able to generate a reference signal identical to the one sent. The incoming signal is then correlated with the reference signal. The use of wideband ultrasound allows for good background noise toleration and sending multiple ultrasound signals at the same time. Accuracy is around 2 cm.

## 2 RELATED RESEARCH AND BACKGROUND

- “Low Cost Indoor Positioning System” (Randell 2001) uses four transmitters placed in each corner of the room under the ceiling. An RF signal is sent first followed by a precisely timed sequence of ultrasound chirps, one from each transmitter. This allows the receiver to calculate the time-of-flight for each chirp using the time the RF signal was received and individual signal delays. Accuracy is around 10-25 cm.

The positioning systems mentioned above determine the time the signal was sent by either sending an RF “trigger” to which a transmitter is expected to respond immediately or by sending an RF notification together with an ultrasound signal. This is connected with several uncertainties: delay between the command to send the signal and the signal being sent, the time it takes to encode the signal, a signal’s time-of-flight, the time it takes to decode the signal and delay between the time the signal was received and the system responding. This is particularly problematic for devices that support multitasking such as modern smartphones.

It was shown by Borriello et. al. that 21 kHz ultrasound signals can be successfully emitted and received with conventional desktop speakers and microphones (on a HP iPAQ 3870 PDA and a Dell Inspiron 8200 laptop) (Borriello 2005). The signal was also successfully detected 100% of the time within a range of 10 meters. This was done using three instances of the Goertzel algorithm: one in the 21 kHz frequency and the other two in adjacent frequencies above and below (Banks 2002). The first instance was checked against the other two in order to distinguish the signal from background noise. In order to check how well the detection system copes with common environmental noise three separate tests were performed. One involved a number of people having a conversation, the second involved playing a variety of music recorded in two different formats (mp3 and ogg), and the final test was leaving the system running in an office

environment for two consecutive days. During the three tests the detection algorithm did not detect any signals. This is a very encouraging finding, because it means that it may be possible to keep working with “raw” sound without introducing complicated filters to check for false positives. The biggest source of false signals appears to be the multipath effect, which we hope can be countered with correct placement of microphones and some adjustments in detection algorithms like those proposed in (Peng 2007).

At the moment ultrasound positioning is the most accurate solution for indoor use. It easily passes the one-meter threshold and comes very close to the one centimetre threshold. So far it has been done with the help of custom hardware, but we see no reason why it could not be done using conventional speakers and microphones. The work reviewed indicates that it is possible to make distance measurements based both on the acoustic signal time-of-arrival on a mobile phone and send/receive ultrasound signals inaudible to the human ear. In a recent study Packi et. al. (Packi 2010) demonstrated the use of conventional speakers to locate an array of four microphones pointing in different directions with the help of ultrasound trilateration.

### 2.1.6. Dead Reckoning

**Dead Reckoning** (DR) is a method of estimating position by continuously projecting current direction and speed over the known initial position (fix). In other words DR attempts to model the path the subject took and use that to derive current location (Bowditch 1995). Apart from the fix, which has to be provided by some other positioning system or input manually, a DR system only needs a constant flow of information about current direction and speed. The main feature of this method is that there is no need for external references and therefore no infrastructure need be placed.

## 2 RELATED RESEARCH AND BACKGROUND

Methods of obtaining speed and direction mostly depend on the applications of the system and what resources are already available. For example, in case of a differential drive on a car or a robot, velocity and direction can be easily deducted from the ground contact speed of each wheel and distance between them. The method is not ideal, as factors such as wheel slippage and uneven ground can introduce some error, but in many situations such as indoor positioning within the boundaries of one floor it can work very well. Unfortunately, mobile devices are a much more restrictive environment for performing Dead Reckoning. The user moves independently of the system and there is no contact with the surface he walks over. Also, there is no way to know how the person will carry it: will he hold it in front of himself; put it on his belt; in his pocket; in his bag. The only way to carry out DR under such conditions is to assume the functionality of an Inertial Navigation System (INS). An INS is able to acquire all necessary input without coming in contact with any surface or signal source, and instead uses natural phenomenon such as gravity or Earth's magnetic field (Siciliano 2008).

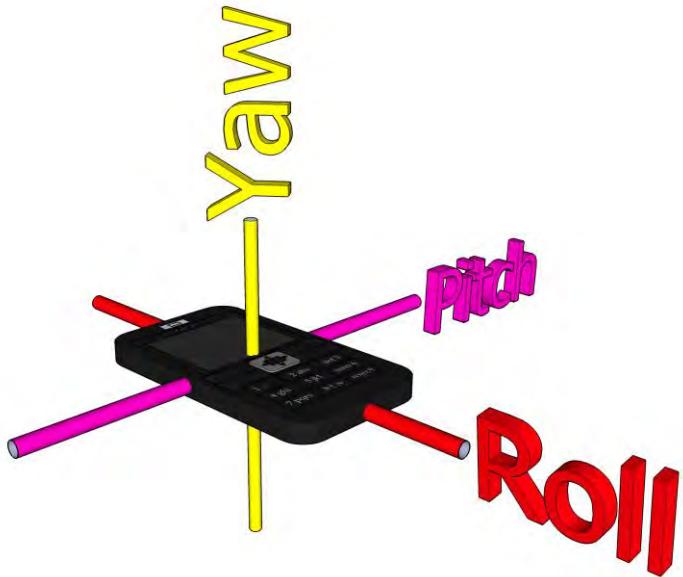


Figure 9: Roll, Yaw and Pitch axis.

## 2 RELATED RESEARCH AND BACKGROUND

It is possible to determine orientation of a mobile phone if the following angular/spatial variables are gathered in real time: pitch angle, yaw angle and x,y,z coordinates. Pitch is an angle of rotation in the vertical plane (i.e. an angle in the up and down direction) and can be measured either from the Zenith (up) position downwards or from the Nadir (down) position upwards. (Figure 9)

A **gyroscope** can provide all three angles. The original gyroscope dating back to 1850 relied on the principle of the conservation of angular momentum by suspending a rotating wheel in such a way that it is free to change its axis of rotation. As the device changes orientation, the wheel will remain rotating in the same plane and the resulting difference can be measured. This method has a lot of problems, such as being bulky, difficult maintenance, and is mostly used for demonstration purposes. Modern gyroscopes exploit rotating frames of reference that show specific physical properties, which are measured. For example an optical gyroscope uses the Sagnac effect (Anderson 1994). When two laser pulses are emitted in a fibre-optic ring in opposite directions, they will arrive to the starting point simultaneously. However when the ring is rotated around the centre of the plane it lies in, one of the pulses will take longer to arrive than the other. If this difference is measured, it is possible to measure the rate at which the ring is rotated. Because such gyroscopes measure rotation only in one plane it is common to install them in sets of three (Siciliano 2008).

All modern gyroscopes have one common flaw: gyroscope drift. The problem arises from the fact that the current angle of rotation is calculated by constantly measuring the rate of rotation. Each measurement contains some small error, possibly only a fraction of required accuracy, which accumulates over time. Unless the error is corrected through reference to some other method, the error will eventually exceed the limit of required accuracy. An alternative approach is to use accelerometers to measure Pitch

## 2 RELATED RESEARCH AND BACKGROUND

and Roll angles, plus a digital compass (magnetometer) to measure Yaw angle (Randell 2003; Siciliano 2008). In comparison, accelerometer readings on an individual basis are a lot noisier than readings produced by gyroscopes<sup>17</sup>. Gyroscopes are not found in devices such as mobile phones very often. Accelerometers, however, are becoming ever more popular, being used for example to automatically switch between portrait and landscape screen views on many smartphones currently available.

**Accelerometers** can be used in two ways. First of all they can measure the acceleration of a device (i.e., increase/decrease in speed) in one or several directions simultaneously. Secondly because they are also sensitive to the direction of gravity, it is possible to measure rotation with respect to this direction. When an accelerometer is at rest it registers 1G force along the vertical axis. To acquire acceleration only from movement, local gravity has only to be subtracted. When pitch or roll angles of a phone change, the gravity vector starts to point in a different direction with respect to the accelerometer, which can be measured. To understand why the yaw (horizontal) angle cannot be calculated with an accelerometer, consider a plumb line attached to the middle of the phone. When the vertical angle between the plumb line and the phone's surface changes – that can be measured. But when yaw angle changes, it just means the phone has rotated about the plumb line.

A **Digital Compass** or Magnetometer is a device that can determine the direction of the Earth's magnetic field. Not unlike a traditional compass, it can be used to determine orientation in the yaw axis. A magnetometer usually shows direction in relation to

---

<sup>17</sup> "Invensense: Video Library" Accessed 12 November, 2012, at  
<http://invensense.com/mems/videolibrary.html>

## 2 RELATED RESEARCH AND BACKGROUND

magnetic north, so an offset has to be applied according to the geographical location in order to reference true north. The biggest problem for indoor use is interference from various localised magnetic fields. Things such as electronic equipment, magnets (e.g. that hold doors) and large iron objects can locally distort the Earth's magnetic field and make compass readings unusable. Also, experiments show that even outdoors where no significant interference was observed, a digital compass produces noisier results than a gyroscope (excluding drift) (Randell 2003). Unfortunately at the moment there is no better way to determine yaw orientation on a mobile phone, which is a lot more important for DR than Roll and Pitch. We think that one possible solution could be to carry out fingerprinting of magnetic distortions similarly to how it is done for GSM positioning (see page 21). Some advances have been made recently in utilising Kalman filters for this task (Goyal 2011).

There are several ways to estimate current velocity using accelerometers. The most straightforward approach is to monitor acceleration or deceleration of the unit, which is exactly what an accelerometer does. In an ideal scenario, where vibrations do not exist and all external forces (except gravity) contribute to movement, this approach works very well. Unfortunately, accelerometers are very sensitive to vibrations, which is a major problem for mobile devices. This is probably the main reason why this method is almost exclusively used for more stable robots and vehicles. From the literature, there are currently no reliable implementations of this method for pedestrian use. Another problem is that, similar to gyroscopes, accelerometers register the rate of change in a user's movement speed, which leads to build up of positioning error over time (Siciliano 2008).

A much more reliable use of accelerometers for estimating pedestrian movement is step detection. When we walk, every step is associated with upward and downward motion

## 2 RELATED RESEARCH AND BACKGROUND

of the torso. This movement is harder for accelerometers to pick up than the movement of feet, but it is still possible. The spikes of acceleration are used to detect if and how frequently the user takes a step. In a study done by Randell et. al., two factors were used to help establish speed (Randell 2003). When humans stride further, the legs stretch out further coupled with feet moving faster to not lose stability. Therefore the vertical acceleration of the foot is greater when the foot makes a larger stride. Another observation is that humans rarely under-step their minimum average step unless turning on one spot. Of course this minimum average step is different for everyone. The tests were done with different configurations of equipment and the overall conclusion was that provided the user received some sort of training, pedestrian Dead Reckoning can provide better accuracy than GPS over short distances. However if an untrained user was equipped with such system, the results would be worse. This can be addressed to some extent by calculating the individual user's step during runtime (Goyal 2011).

There are a number of additional problems connected with implementation of DR on a mobile device. By nature, Dead Reckoning has to work continuously and every DR implementation expects the INS to be at a fixed or controlled position and orientation in relation to the vehicle or in our case a person. This contradicts with our everyday use of mobile phones. When we work with a phone, it is held in one hand in front of our eyes or in two hands in a landscape mode. When we do not work with it, it could be in our pocket, or on the belt or even swinging along with the arm it is held in. If we take for example the case when the user holds the phone in his hand and swings his arm as he walks, the readings from accelerometers will be completely useless. While the above sounds only somewhat likely, a much more common problem is lack of knowledge about the orientation of the phone in relation to the user. A step-based approach only gives us the speed and assumes the direction is the same as the yaw angle value provided by a gyroscope or a magnetometer. Unfortunately that is not the case. If the

user is not working with the phone right now, the phone could be facing any direction depending on how the user prefers to carry it. Finally poor performance of magnetometers in indoor environments is a major issue that will have to be resolved in order to carry out directional querying.

On the whole, readings provided by accelerometers and a magnetometer should be used to improve accuracy and enhance functionality of other positioning methods. From this point of view their presence in modern smartphones is very valuable. However an indoor positioning system that uses Dead Reckoning implemented on a mobile phone will have to be operated very carefully or will be completely unreliable.

### **2.1.7. Computer Vision Approach**

Computer vision is concerned with computers extracting information from images and video. Among other things, computer vision can be used to determine current position by analyzing a live video stream from a video camera. In principle, computer vision positioning relies on recognising certain visual features that happen to be in the camera's field of view and determining their position and orientation in relation to the camera with the help of known properties of the camera such as its focal length and known size and shape of the objects. The camera can be either mounted on the wall in which case its precise position and orientation will be known and it is possible to deduce position and orientation of the visual features tracked, or alternatively the camera can be a part of a mobile device that needs to be tracked in which case the position and orientation of tracked visual features (reference points) will either have to be known or acquired during runtime. The first method is usually used to track faces and possibly recognise gestures. Tracking a mobile phone would be a very challenging task as it will be often obstructed and hard to recognise. Displaying a fiduciary marker

on the screen may help, although the screen will probably be too small. This method does not take advantage of a mobile phone's hardware and is technically outside the scope of this research. On the other hand tracking with a mobile camera is relevant, because practically every modern smartphone has an inbuilt camera.



**Figure 10: Fiduciary marker.** Used to determine orientation and distance to camera.

Visual reference points or control points can either be artificially added into the environment or already be a part of it. In the first case it is common to use fiduciary markers (see Figure 10). These markers are perfect for determining position and orientation for a number of reasons. The black and white pattern has high contrast and most of the time can be easily extracted even with such simple filters as thresholding. The square frame is ideal for determining size and orientation. Currently there are a few very quick algorithms that can determine these two parameters in real-time even on mobile devices or in Flash (Wagner 2003). The area inside the square frame can be occupied with a unique pattern so that the program can tell one marker from another. If only one marker is used or it is not necessary to distinguish between markers, some sort of image inside the square will be necessary to tell which side of the square is the top. It is common to put a dash at the bottom of the square for this purpose. Some good

## 2 RELATED RESEARCH AND BACKGROUND

examples of the technology are AR Tower Defence by Cellagames<sup>18</sup> and Smart Grid Augmented Reality<sup>19</sup>.

In order to determine position a purely marker-based positioning system requires a unique marker to be in the field of view of the camera. The marker also has to be close enough for the system to correctly estimate distance and angle. This presents a difficult problem of placing the markers in the environment. Probably the best surface for fiduciary markers is the ceiling. We almost never come in contact with the ceiling, thus there is little danger of the markers being damaged. Regardless of where a user stands in the room, he will be at the same distance from the ceiling, which is important to ensure reliability. Patterns on the ceiling are less likely to be blocked by furniture or people. Finally, because we rarely look at the ceiling, placing markers there will have smaller impact on the overall look of the premises. Nakazato et. al. used this approach in combination with a helmet carrying a Head Mounted Display (HMD) and camera pointing upwards (Nakazato 2005). Because either device cannot move in relation to the other, it is possible to convert the positioning data collected by the camera for HMD to provide Augmented Reality (AR). Unfortunately markers on the ceiling are not as suitable for mobile phones. Smartphone cameras are usually located on the back of the phone and very rarely on the front. Normally when the user is working with the phone he will hold it at an angle and the front camera view will be obstructed by the user's head. The camera at the back of the phone will likely have a patch of the floor in its view and probably a lower part of a wall if there is one close by. Very often this zone is occupied by furniture. In addition, the floor is a walking surface, and embedding

---

<sup>18</sup> "Cellagames.com" Accessed 12 November, 2012, at <http://cellagames.com>

<sup>19</sup> "GE: Plug Into the Smart Grid" Accessed 12 November, 2012, at  
[http://ge.ecomagination.com/smartgrid/#/augmented\\_reality](http://ge.ecomagination.com/smartgrid/#/augmented_reality)

## 2 RELATED RESEARCH AND BACKGROUND

markers in it is expensive compared to printing markers on plain paper with a laser printer. In early versions of the Signpost project a handheld device was used rather than HMD and markers were placed on the walls at eye level (Wagner 2003). Because Signpost is based on AR, the user looks through the handheld's screen when he works with it, so the placement of markers at such height makes sense. Also, the Signpost system seems to mainly target corridor navigation, because in a large room markers will often be too far away.

Just like with Dead Reckoning it matters how the user holds the phone, but an important difference is that failing to detect any markers is not going to affect future accuracy in any way; the only thing that will happen is the device will not be able to tell the current position. Logically this means that the markers can be placed only in areas where the user has use for a positioning system. Considering directional querying is an important interface for an indoor positioning system, this could be seen as a potential area for optimisation. For example fiduciary markers could be placed right next to the objects of interest, eliminating the need to tie down the coordinates of all markers, walls and large objects into one coordinate system. In fact the system could be simplified even further by placing the markers on the objects of interest (if possible) thus eliminating the need to update their location when they are moved. The above two approaches are not positioning systems anymore and therefore deviate from our original objective. Nevertheless it should be noted that such an approach is a good alternative to a full-blown positioning system if all that is necessary is to display information about an object.

The use of fiduciary markers introduces one aesthetical issue: the markers are easily visible to the human eye and very rarely fit into indoor design. At first glance this is unavoidable, because digital cameras operate similarly to human eyes, and naturally a

## 2 RELATED RESEARCH AND BACKGROUND

pattern on a marker has to be as vivid as possible for computer vision to work better. However digital cameras see a broader spectrum of light and it may be possible to exploit that. For example Kameraflage<sup>20</sup> allows the production of cloths that appear to have a bright blue pattern/text on black background when viewed through a camera while to human eyes the same area appears to be just black. It appears to be necessary for the background to be black, which indoors is rarely found in necessary abundance, plus it has to be well illuminated.

It was suggested by Nakazato et. al. to use retro-reflectors (Nakazato 2005). A marker made with retro-reflectors is white and under normal conditions appears to have a less than prominent pattern to a human eye. When illuminated with infrared LEDs and captured by an infrared camera, the marker will appear as a white distinct pattern on a black background. Smartphones, unfortunately, do not have infrared cameras and under normal lighting retro-reflectors will appear almost imperceptible to standard cameras. Remarkably, the pattern will be visible if captured on a standard camera with a flash (see Figure 11).



**Figure 11: Retro reflectors captured without and with a flash, adapted from Nakazato Y. (Nakazato 2005).**

---

<sup>20</sup> "Kameraflage.com" Accessed 12 November, 2012, at <http://www.kameraflage.com>

It is possible to determine pose of a surface without a marker if the surface has at least some kind of unique pattern. With the help of an algorithm such as Scale-Invariant Feature Transform (SIFT) it is possible most of the time to recognise a 2D shape regardless of its size, position, and orientation in 3D space (Lowe 1999). Initially, interesting points have to be extracted from the original 2D shape in order to get a “feature description” of the object. Later during runtime the SIFT algorithm will analyze every frame of the video feed and recognise the shape based on its “feature description”. After this the pose of the surface can be established based on how 3-4 SIFT points in the original shape and the video frame are related by homography. An important advantage of this approach is that it will still work if there is a partial occlusion. Experimentally it was shown that SIFT-based positioning is possible outdoors during daylight by using building facades as reference surfaces (Bres 2009). Unfortunately it was not possible to make it work in real time even on a laptop as the SIFT detector is computationally heavy. Also, unlike fiduciary marker that can tolerate rotation of up to 90°, most invariant point descriptors work only at 40-50° of tilt.

Because SIFT is so computationally intensive, it was proposed to use Features from Accelerated Segment Test (FAST) corner detector to detect features (Wagner 2008). While it is significantly faster, the descriptor is no longer scale invariant. To reintroduce scale estimation, the descriptor database contains features from all meaningful scales. Consequently, this is trading memory for speed (Wagner 2008). Depending on the specific phone configuration, this approach may turn out to be more reasonable for mobile devices.

FAST can be used to detect and track features in real time without a training phase in what is call Simultaneous Localization and Mapping (Williams 2007). This technique is

## 2 RELATED RESEARCH AND BACKGROUND

extensively used in robotics to map an unknown environment. The procedure is very complicated and involves constantly rewriting or updating a map of the environment while at the same time localising the camera's position in that map. While traditionally it has been used for robots, where it is possible for the system to control the camera's movement as well as obtain non-visual information about the robot's movement, there are examples of using SLAM successfully with a camera operated by a user (Siciliano 2008). It would be interesting to discover whether SLAM can run sufficiently fast on a mobile phone. In the experiments done by Williams et. al. it took on average 19 ms out of the budget of 33 ms for one cycle of their SLAM implementation to complete on a computer with a Core 2 Duo 2.7GHz processor (Williams 2007). There is no doubt some frame rate will have to be sacrificed in order for SLAM to run on a mobile platform, but how badly this will influence usability is unknown.

In terms of directional querying the biggest advantage of computer vision is that it delivers orientation as well as position. This is particularly important indoors because, as it has already been noted, magnetometers perform poorly in such conditions. Another advantage is that infrastructure can be very cheap or there is no need for it at all. In terms of accuracy it is hard to evaluate computer vision methods because it depends on the number and distance to visual clues in every single frame. On the other hand, considering the object of interest is visible within the frame and its size is proportional to the distance to it, the notion of accuracy that was used to describe other non-visual positioning systems is no longer applicable. This inevitably leads us to an observation that most computer vision systems are either designed to be autonomous (e.g. a robot) or communicate via Augmented Reality. Performance is probably the weakest aspect of computer vision systems on mobile devices, which is why many developers keep coming back to the idea of outsourcing the computational load to a server (Wagner 2003; Wagner 2008).

Other than positioning, it has also been proposed to use live camera feed analysis to detect changes in phone's orientation(Wang 2006) on devices that lack accelerometers and for hand gestures(Kratz 2007). More recently it was proposed to use it for step detection(Aubeck 2011) in order to improve accuracy of Dead Reckoning systems or to track changes in the phone's orientation (Ruotsalainen 2011) as an alternative to magnetometers.

Overall computer vision appears to be a very promising indoor positioning method. There is a great variety of techniques to choose from. It can be as simple as assigning every object of interest a unique fiduciary marker or as complicated as SLAM. In particular our expectations lie with SLAM, as many people recognise it to be a very promising technology still in its infancy (Siciliano 2008).

### 2.1.8. Discussion

A comparison of positioning methods discussed in this thesis can be found in Table 1.

**Table 1: Comparison of positioning methods for smartphones.**

*Rows correspond to positioning methods and columns correspond to parameters.*

	<b>works indoor</b>	<b>relative accuracy</b>	<b>infrastructure cost</b>	<b>reliability</b>	<b>performance</b>
<b>GPS/AGPS</b>	no	poor (n/a)	none	good	good
<b>GSM</b>	yes	average	none	good	good
<b>Wi-Fi</b>	yes	good	none/average	good	good
<b>Bluetooth</b>	yes	good	average	good	good
<b>Sound</b>	yes	excellent	average/expensive	good	good
<b>Dead Reckoning</b>	yes	degrades	none	poor	average
<b>Computer Vision</b>	yes	excellent	none/average	poor	poor

## 2 RELATED RESEARCH AND BACKGROUND

It is hard to compare accuracy of positioning methods in general, because the numbers greatly vary between implementations. What makes this even harder is that the numbers for one implementation may also vary greatly between different environments in which experiments were carried out. Therefore if we were to compare only the best results for the most accurate implementations, true merits and demerits of the methods would not be reflected, because in some cases a particularly good result can be achieved only under very specific conditions. Similarly stating the entire accuracy range is not very effective either. For example, closest known location for GSM in a remote suburban area will yield a radius of several kilometres. Consecutively using an average is not going to work well either. Instead we attempt to compare how well each of the methods would fare against each other under the same indoor conditions based on actual limitations of each of the underlying technologies and backed up by experimental data rather than just based on numeric data. Best accuracy results of all implementations mentioned in this paper are given in Table 2 for reference purpose.

From the literature it appears that GSM positioning on average has twice worse accuracy than Wi-Fi positioning. Bluetooth positioning has accuracy very close to that of Wi-Fi. Compared to the above methods, sound positioning is many times more accurate. Because Dead Reckoning does not really work on portable devices, it is probably pointless to estimate its accuracy. While it may potentially deliver good accuracy initially, it will quickly degrade and, if used carelessly, very quickly fail. Accuracy of computer vision approach even within one implementation and experiment depends on a number of factors including distance to the nearest marker/feature, lighting and motion blur. Potentially it can be very accurate, but what is more important is that its reliability once again depends on user's actions. This is an important factor, not just because the user has to be instructed or even trained, but also because it makes the user's experience with the system more bothersome, forcing the user to actively seek

## 2 RELATED RESEARCH AND BACKGROUND

out markers or keeping the phone at a particular tilt angle and as a result decreasing the value of such service. The concerns with the likeliness of each method to fail or degrade unexpectedly are reflected in the “reliability” column.

**Table 2: Comparison of indoor positioning implementations.**

*Rows correspond to positioning implementations and columns correspond to parameters.*

	best accuracy	underlying technology	available on smartphones
<b>Wide signal strength fingerprinting</b>	2.48m	GSM	no
<b>Skyhook(GSM)</b>	200m	GSM	yes
<b>Navizon(GSM)</b>	50m	GSM	yes
<b>Skyhook(Wi-Fi)</b>	10m	Wi-Fi	yes
<b>Navizon(Wi-Fi)</b>	20m	Wi-Fi	yes
<b>Hybrid Fingerprinting</b>	3m	Wi-Fi	yes
<b>RADAR</b>	2m	WaveLan	no
<b>GP for Signal Strength-Based Location Estimation</b>	2m	Wi-Fi	yes
<b>Goodtry</b>	4m	Wi-Fi	no
<b>Ekahau</b>	1m	Wi-Fi	no
<b>Bluetooth Direction of Arrival</b>	72cm	Bluetooth	no
<b>Nokia High Accuracy Indoor Positioning</b>	20cm	Bluetooth 4.0	not yet
<b>The Bat</b>	3cm	Ultrasound	no
<b>The Cricket</b>	3cm	Ultrasound	no
<b>DOLPHIN</b>	15cm	Ultrasound	no
<b>High Performance Privacy Oriented location system</b>	2cm	Ultrasound	no
<b>Low Cost Indoor Positioning System</b>	10cm	Ultrasound	no
<b>SNoW Bat</b>	15mm	Ultrasound	no

Another parameter that should not be underestimated is infrastructure cost. Some methods ultimately need no additional infrastructure, such as GSM. For methods such as Computer Vision infrastructure can vary a lot: there may be need for a server that the

phone could outsource the computational load to, or it could be as simple as paper markers printed with a black and white laser printer or even none at all. The cost of infrastructure for sound positioning is marked as potentially expensive, because hardware that allows simultaneous processing of data from several microphones is inconveniently expensive. There is probably a way to avoid using professional recording equipment and we see this as one of the more important challenges of sound positioning.

Finally the “performance” column was added to reflect the computational load on a mobile phone, which is characteristic to most computer vision methods. The necessity of DR to run uninterrupted and address sensors several times per second was also reflected here.

The review of positioning technologies can be summarised in a few observations:

- It is currently impossible to achieve accuracy below one meter threshold using electromagnetic signal propagation and conventional mobile phone hardware.
- Positioning based on sound propagation can deliver accuracy up to several centimetres.
- Methods that are not based on signal propagation are usually tricky to operate especially for an untrained user.

## 2.2. Services

Location based services have already entered our everyday life. Maybe not yet for everyone, and maybe they are not as useful or widespread as it has been predicted many times in recent history, but the first step was made years ago and as mentioned in the

## 2 RELATED RESEARCH AND BACKGROUND

introduction, there were a couple of conditions for LBS to emerge. For example, although GPS has been largely available for some time, it was enabled via standalone devices that did not allow you to do anything except navigate from one location to another. It was not until GPS receivers started to appear inside mobile phones that things changed. Two distinctly different examples can be given. One is the location-augmented iMode website marketed by NTT DoCoMo<sup>21</sup> in Japan which was used for dating among other purposes. Another example is an old and simple program for Windows Mobile that would let you remember certain places and attach reminders to them (Kolodziej 2006). It was useful, for example, if you needed to collect a flash key from a friend but it was not too urgent and you did not know the next time you will be in the area. LBS may or may not need information about the environment or other users, but what it absolutely needs is the means to determine position and a platform to execute various programs on. Simple GPS navigation systems lack the second element and it comes as no surprise that the appearance of fully-functional GPS navigation for free on smartphones, such as Google Maps Navigation puts manufacturers of dedicated navigation systems at a serious disadvantage (Arrington 2009).

Currently there are no examples of fully-functional indoor LBS for mobile phones, but theoretically they could perform a number of functions:

- Make evacuation procedure more intuitive and efficient by showing directions along the shortest path (Meijers 2005). In this example it is important for the system to know 100% of the time where the user is so that he does not have a reason to panic if he suddenly realises he got lost.

---

<sup>21</sup> "NTT DOCOMO" Accessed 12 November, 2012, at <http://www.nttdocomo.com/>

## 2 RELATED RESEARCH AND BACKGROUND

- Improve navigation in shopping malls. There is already a company that collects and maintains maps of shopping malls<sup>22</sup>. Normally when working with an unfamiliar map it takes a significant amount of time to figure out current position and direction unless the map is stationary and the position is already marked. This makes portable maps less useful. Using indoor positioning it is possible to take better advantage of such data. Showing the current position on an interactive map would already be a significant improvement and giving instructions how to get to a particular shop would make navigation even easier.
- Given better precision it may be possible to direct the user to a particular shelf in a shop<sup>23</sup>. Bearing that in mind it is possible to design a program where the user has been populating a list of things he needs to buy on his mobile phone since he last went shopping. When he enters a shop the most optimal route to collect the goods is generated and the user is instructed where to go next.
- A library catalogue (Bahl 2000) combined with a navigation system that directs the user to the shelf with the book he requested.
- A museum virtual tour guide (Chou 2004; Tsai 2010). Systems currently used in museums provide very unsophisticated functionality which is very often limited to pointing at a tag or manually entering a number in order to hear a recording. A system with true indoor positioning based on a mobile phone can be used by pointing at the actual exhibit and not at a tag via directional querying. Depending on the arrangement and size of exhibits, directional querying may require very high spatial and directional accuracy. A smartphone can deliver a variety of content including audio, video, text, images or a combination of them such as a webpage.

---

<sup>22</sup> "Point Inside" Accessed 12 November, 2012, at <http://www.pointinside.com>

<sup>23</sup> "Nearbuy" Accessed 12 November, 2012, at <http://www.nearbuysystems.com/>

## 2 RELATED RESEARCH AND BACKGROUND

Once again because the system is continuously aware of user's location it is possible to guide the user to an exhibit he wants to see, to the exit or any other facility.

- Use in a company to track employees. Systems currently used for this purpose use Wi-Fi or RFID tags<sup>24</sup>. The main problem with using tags is that while the person controlling the system knows where everyone is, an average user has no benefit from this system. A smartphone version however can offer any employee to find any other employee regardless if he is right now at his desk or not. Depending on the type of work this ability may turn out to be extremely valuable. Also it is not unusual for companies to issue smartphones such as Blackberries to every employee, so it is very likely that everyone is already carrying necessary hardware.
- There are a lot of more specialised fields such as equipment for warehouses; however a comparison to professional tools already used in these fields is beyond the scope of this research.

An important characteristic is that the services mentioned above are not just capable of running on the same device but are not necessarily different software. For example the evacuation system can and should be implemented in every other system on the list, considering all the necessary assets such as the model of the premises and positioning infrastructure are already there.

---

<sup>24</sup> "The Best of Both Worlds." Retrieved 10 June 2012, from  
<http://www.intelleflex.com/downloads/white-papers/Best-of-Both-Worlds.pdf>.

### 2.3. Interfaces

The most common interface to deal with spatial data in LBS on mobile devices is an interactive map. There are a few applications that try to take advantage of the mobile nature of the device. It is interesting to note that these interfaces largely intersect with the list of next generation Geographic Information Systems (GIS) Egenhofer proposed in 1999 (Egenhofer 1999).

**Geo Sketch Pads** offer a multi-modal interface where the user can write notes on top of pictures he has taken and attach coordinates and direction. This creates a connection between data captured by the recording device and the person's feelings at that moment. These features recently became available in digital cameras. There are a few models that have GPS receivers built in. They attach coordinates to the pictures at the moment they are taken, while photo hosting services such as Flickr have learned to read this data and overlay their coordinates on a map. A lot of cameras also allow audio comments to be recorded and attached to pictures.

**Smart Compasses** display an arrow on the screen which points in the direction of user's destination. There are very few GPS navigators available now that have a digital compass (i.e. magnetometer), so in other navigators this service is only available at the speed of a car where direction can be inferred by composing multiple coordinate readings over time. On smartphones that have magnetometers this service can be a welcome addition for pedestrian navigation. In some situations it is better to just see direction rather than read a map. Also some people simply do not want to deal with maps, like while riding a bike. Recently phones started to rotate the map depending on which direction you are moving so your heading is always pointing "up", which could be considered an extension of the smart compass idea.

## 2 RELATED RESEARCH AND BACKGROUND

**Smart Horizons** allow a user to look beyond his/her field of view. When he points his phone in a particular direction, depending on the purpose of the particular program, certain information will be displayed about that direction. It could be buildings, traffic, weather. The idea is to stop the horizon hampering the user's decisions. Technically a lot of applications offer the ability to view information about a certain place; however the gesture interface is missing, which is an important part of the implementation. Instead some applications such as project Enkin<sup>25</sup>, Layar<sup>26</sup> and Wikitude<sup>27</sup> use Augmented Reality (AR) to superimpose tags with information on the live feed from the phone's camera. The user can choose the radius of virtual horizon which will determine how close an object has to be for its tag to be displayed. Just like in the interface proposed by Egenhofer (Egenhofer 1999), this should help the user disregard constraints imposed by horizon or nearby buildings, but the implications are a bit different. We believe that the method presented in these three applications does not really stand up to the idea of helping a person look beyond his field of view. Information given about a remote location is limited to roughly which direction it is and how far it is, but does not give a very good perspective of the surrounding area or position in relation to other points of interest. Perhaps a specialised application made for professional use, for example engineers, pilots, or foresters would better illustrate the benefits of this approach. Finally there is a program called Street View<sup>28</sup> which integrates 360 degree pictures taken on street level into Google Earth<sup>29</sup> and Google

---

<sup>25</sup> "Enkin Blog" Accessed 12 November, 2012, at <http://enkinblog.blogspot.ie/>

<sup>26</sup> "Layar" Accessed 12 November, 2012, at <http://layer.com/>

<sup>27</sup> "Wikitude" Accessed 12 November, 2012, at <http://www.wikitude.org/>

<sup>28</sup> "Street View" Accessed 12 November, 2012, at <http://maps.google.com/intl/en/help/maps/streetview/>

<sup>29</sup> "Google Earth" Accessed 12 November, 2012, at <http://www.google.com/earth/index.html>

## 2 RELATED RESEARCH AND BACKGROUND

Maps<sup>30</sup>. On Android phones it has a feature which allows scrolling through a 360 degree picture by rotating the phone around you. This works thanks to the built in magnetometer. Unfortunately there is no correlation between the direction the phone is pointing and the view of the picture other than the same spatial orientation. So for example we are in Dublin and can browse a picture of a street in Paris. While the phone points north, the viewport in the picture will also point north, however the direction your mobile phone is pointing does not have anything to do with where the location of the street is. If such correlation was introduced, the technique may actually qualify as a Smart Horizon.

**Geo-Wands** allow users to identify objects by pointing at them. This should replace the traditional use of map and compass. The M3I platform has a very similar feature (Wasinger 2003) that allows the user to use speech and gestures to allow a more natural interface. Also they support two kinds of gestures: *intra* – when a user points at something on the screen and *extra* – where he points at a real object using the phone as a pointing stick. There are relatively few applications that allow such *extra* gestures. While *intra* gestures are completely different from what Egenhofer has described, they are used for the same purpose in multiple mobile applications. Just like with today's implementations of the smart-horizon idea described in the previous paragraph, superimposing messages on top of the live video feed from the phone's camera is involved. It is difficult to draw a line between the two. Technically, smart-horizon applications are supposed to work with objects you cannot see at the moment, which means you cannot really point at and identify them. On the other hand there is nothing stopping a user from using smart-horizon software to identify a building right in front of

---

<sup>30</sup> "Google Maps" Accessed 12 November, 2012, at <https://maps.google.com/>

## 2 RELATED RESEARCH AND BACKGROUND

them, in which case the program will really act like a geo-wand. Therefore we will classify applications as geo-wands if they are either unable to identify objects that are hidden from your field of view or such restrictions have been introduced for design reasons. The work done by Bres and Tellez falls into the first group because it utilises computer vision and pattern recognition (Bres 2009). While the Mobile Application Framework for the Geospatial Web done by Simon and Fröhlich falls into the second group, because their application interface is designed to look like the user's line-of-sight, so that buildings hidden from the user's view do not appear on the display (Simon 2007). Geo-Wand type interfaces are probably going to benefit the most from positioning on mobile phones with sub-meter accuracy. Given that many smartphones have hardware that can be used to determine the phone's orientation, with some effort it should be possible to enable directional querying of relatively small objects.

**Smart Glasses** are a lightweight Head Mounted Display (HMD) that use Augmented Reality to display relevant information in front of the user's eyes. This is the exact same AR described earlier (Section 2.1.7), but there is a difference in usability requirements. Because the display is everything the user sees while the glasses are on, the superimposed graphics should not impede spatial orientation. AR will usually display labels near the objects of interest or in more advanced cases display 3D objects as if they were a part of the scene. Either way the system needs to be aware of what part of the scene is in the viewport. It makes sense to implement Augmented Reality on HMD only if labels are directly attached to objects of interest and 3D models are superimposed as if they were a part of the environment. Therefore a lot more accuracy is required. If inaccurate measurements result in the 3D object being tilted, or positioned at the wrong distance, or incorrectly clipped because only a part of it is visible, it will result in the "suspension of disbelief" being ruined. This means that information about

position should be directly extracted from the camera feed via machine vision rather than acquired indirectly from other sensors.

As this solution requires hardware not currently found on a smartphone, mainstream applications are not available in this field yet. But there are a number of interesting prototypes. A Japanese user localisation system for wearable augmented reality that uses a combination of invisible markers and infrared camera is of particular interest to us (Nakazato 2005). It works indoors and unlike many other systems that utilise computer vision, it neither needs black and white markers on the walls or a complicated algorithm to build the environment map from scratch. Translucent retro-reflective markers are placed on the ceiling in the form of a dense grid. These markers are white and when observed closely it is possible to spot a pattern on top of them. In infrared however this pattern is very vivid. This is an acceptable modification for most offices, hospitals and probably some museums. The user then wears a HMD with an Infrared (IR) camera and lights that point at the ceiling. IR Lights illuminate the markers and the camera can read the patterns. This allows for estimating a user's position and orientation very accurately and was demonstrated that this system can realistically superimpose labels and directions on walls. The recently announced project Glass by Google is a perfect example of Smart Glasses interface (Rivington 2012).

## 2.4. Summary

Outdoor LBS has become a huge success mainly because it provides an ideal environment for innovation. Currently it has the advantage of reliable positioning via GPS (also Wi-Fi and GSM) and a defined business model for the delivery of content to the user. Also since business owners are interested in their details being present in programs such as Yelp, there is a well-defined content generation model. With that

## 2 RELATED RESEARCH AND BACKGROUND

granted, it becomes significantly easier to concentrate on designing and developing an actual service.

Currently the same cannot be said about indoor LBS. Probably because both content (spatial database) and infrastructure have to be provided by the owner of the building, indoor positioning and context-sensitive services have mostly stayed away from each other. For example, if a company decides to invest into installing a positioning system in their offices, the most likely purpose for it is to track inventory or employees. Tags, such as Wi-Fi tags, are perfect for that. While such a system easily manages its primary task, it is insufficient to deliver a good Location Based Service. On the other hand indoor context-sensitive systems such as museum virtual guides have managed so far without proper positioning. Typically a device is given to the user that can play back comments for particular areas or even particular exhibits. Sometimes the user has to enter numbers manually, sometimes point at a tag. If the switch between areas/exhibits happens automatically without necessarily pointing at something, this is indistinguishable from an LBS. Overall these systems manage at their primary task relatively well. The reason they do not use positioning is that a lot of these devices have to be manufactured and are often lost, broken or taken away by users, so more expensive hardware required for positioning is just not cost effective. What these systems cannot do, for example, is tell the user how to get from where he is right now to the exit. Based on what outdoor LBS has demonstrated, there is a great number of context-aware services that a device equipped with a screen, sensors and a proper OS can provide once the means to establish position are there. This is the primary reason we covered only positioning technologies that are available on modern smartphones.

## 2 RELATED RESEARCH AND BACKGROUND

After having reviewed positioning methods available on most modern smartphones, ultrasound trilateration was chosen as a primary research topic of this thesis for the following reasons:

1. Among positioning methods reviewed, only sound positioning can potentially offer consistent sub-meter accuracy. There are good reasons to aim for higher accuracy of estimated position and orientation. To begin with, everything indoors happens on a smaller scale. Corridors are narrower than streets and room entrances are smaller than shop fronts. An indoor LBS is very easy to expand in terms of functionality once all the infrastructure and spatial data is there, so if there is no need for sub-meter accuracy initially, lack of it should not be a limiting factor for expansion. The requirements for accuracy can be different depending on the task. For example a virtual tour guide with spatial querying will require as fine accuracy as possible at least below one meter, because the deviation will increase as the distance to the object increases. While privacy is a good reason to limit maximum positioning accuracy for pervasive technologies such as GPS, GSM and possibly Wi-Fi, it should not be of concern for sound positioning as it cannot be used to determine position outside the areas equipped with the infrastructure.
2. Ultrasound trilateration is sufficient on its own and will not benefit much from merging with other positioning methods. Among GPS solutions only pseudolites work indoors, but they are currently not compatible with mobile phones. GSM provides no benefit, being less accurate. Some simple form of Wi-Fi or Bluetooth positioning may be used to track the user between locations for extra reliability considering a connection will be needed anyway to send requests and content, however this is not a major issue. Dead Reckoning appears to be practically unusable on mobile devices. Finally computer vision is a very promising solution on its own, but there is little benefit from combining it with sound trilateration

## 2 RELATED RESEARCH AND BACKGROUND

except for maybe virtual tagging of assets in AR applications. While computer vision can be very accurate, it will consume a lot of computational resources; require a lot of development and tweaking while at the same time be dependent on how the user operates the phone.

3. Ability to use ultrasound, which is inaudible to human ears, is an important attribute of a system that uses sound waves. If a sound signal used for trilateration was within the hearing range, it would appear sharp, loud and overall unpleasant to human ear. This is because a signal needs to be as distinct as possible in order to cover long distances, resist reverberation and clearly identify time-of-arrival. The concept is very similar to how fiduciary markers in computer vision must be very vivid to allow accurate readings unless the system uses infrared, which is invisible to human eyes.
4. Sound presents an effective way of using trilateration with conventional mobile phone hardware. Because under the same temperature conditions sound travels through air at a constant relatively slow speed, it is possible to accurately deduce distance from time-of-arrival even at an average sample rate. In contrast, electromagnetic waves travel at the speed of light, so Wi-Fi, Bluetooth and GSM trilateration has to rely on signal strength, which is a much less reliable parameter.
5. Ultrasound positioning is compatible with many mobile interfaces. Because ultrasound positioning will work regardless of how the user holds the device, it is not restricted to a couple of interfaces such as is the case with computer vision (Smart Horizons and Smart Glasses). At the same time high accuracy of positioning means interfaces such as Geo-wand (directional querying) can be used with centimetre precision. Finally ultrasound should not disrupt audio interfaces.

Sound certainly has a number of problems such as reverberation which will have to be addressed. An important point is that this is technically possible with the right

## 2 RELATED RESEARCH AND BACKGROUND

algorithms, settings and arrangement. Conversely, methods based on electromagnetic waves are limited by hardware constraints and physical properties of the signal, which leaves very little room for improvement at least in terms of accuracy. To the best of our knowledge ultrasound trilateration of a mobile phone has not been attempted in other work, and thus represents a novel direction of research.

Information collected and analyzed in the literature review has allowed us to expand on research questions that were proposed in the end of introduction:

**RQ 1: Can ultrasound be reliably reproduced by mobile devices?** The small range of ultrasound available under the standard 44.1 kHz sampling rate may not necessarily be reproduced by the speakers. Because this range is of no significance for a majority of buyers, the speakers could have been manufactured to produce frequencies only up to 20 kHz or even 17 kHz.

**RQ 2: What are the desirable characteristics of the emitted signal?** There are a number of signal properties to experiment with such as volume, frequency, length and shape (e.g. linear increase/decrease of amplitude).

**RQ 3: What is the maximum distance at which an ultrasound signal emitted by a mobile phone can be reliably detected with a microphone?** Sound signals tend to fade with distance and even more so high-frequency signals. At the same time if a signal is very loud it may get distorted by the microphone as well as be audible to some people. Therefore an optimal volume must be found and the maximum distance at which the system can reliably tell it from background noise will be the maximum detection range (functional area).

**RQ 4: Can ultrasound positioning be done asynchronously?** Synchronisation between the phone and the positioning system presents a lot of problems such as clock drift and computationally intensive code on the phone's side. It is desirable to avoid synchronisation by using Time Difference of Arrival.

**RQ 5: What accuracy can mobile asynchronous ultrasound trilateration offer?** An average positioning accuracy will be determined first for a scenario where the microphones and the phone are placed in the same plane. After that average accuracy will be calculated for a scenario where the phone is on a different plane.

**RQ 6: What impact can orientation of the speaker and the way user stands have on accuracy and reliability?** Ultrasound is characterised by being highly directional and having poor obstacle penetration. It is therefore necessary to test how various ways a user can stand and hold the phone affect accuracy and reliability.

**RQ 7: Can background noise cause false positives and how this can be countered?** There is a possibility that some electric device (e.g. router, switch, air conditioner, power adapter etc.) in the room produces sound of the same frequency as the signal used by the positioning system and therefore regularly or irregularly causes the system to “detect” a false signal.

### 3. PRELIMINARY EXPERIMENTS

As a result of the review of existing indoor positioning methods, ultrasound trilateration was identified as a viable and novel approach that can work with off-the-shelf mobile phones. In relation to existing work in the field, our research picks up certain features demonstrated by Walrus (Borriello 2005) and BeepBeep (Peng 2007) projects and attempts to build indoor positioning. In particular these are the use of ultrasound but with normal sound hardware and the use of time-of-arrival to estimate distance on mobile devices. However, before developing a positioning system can be attempted, a number of research questions have to be answered. This is related to the fact that using ultrasound signals produced by conventional speakers to determine time-of-flight has not been attempted before and therefore some aspects of this method must be investigated first.

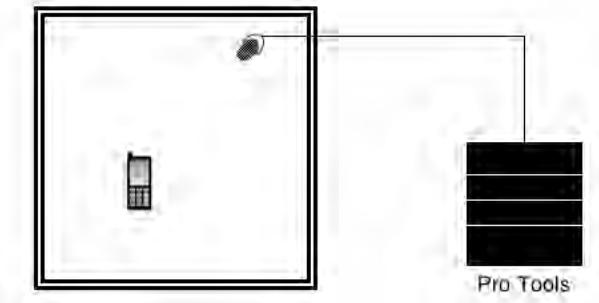
#### 3.1. Ultrasound Generation

This section addresses the first research question (RQ1): Can ultrasound be reliably reproduced by today's mobile devices? The capability to generate and receive ultrasound with conventional sound hardware was demonstrated in the work of Borriello et al. (Borriello 2005), in which the sound was generated with "typical desktop speakers". Whether this will work on any/all modern mobile phones is unreported in the literature and can only be found out through experiments. While a standard 44.1 kHz sampling rate can encode frequencies up to 22 kHz, there is no guarantee that sound hardware will play them correctly. Specifications published by manufacturers usually indicate what frequencies the manufacturer guarantees will work, rather than technical limitations of hardware. Plus such specialised information is much harder to obtain about mobile phones compared to external microphones/speakers.

**Experiment Method.** Experiments were done with four different smartphones:

- HTC G1
- HTC Hero
- Apple iPhone 3GS
- Nokia 6210 Navigator.

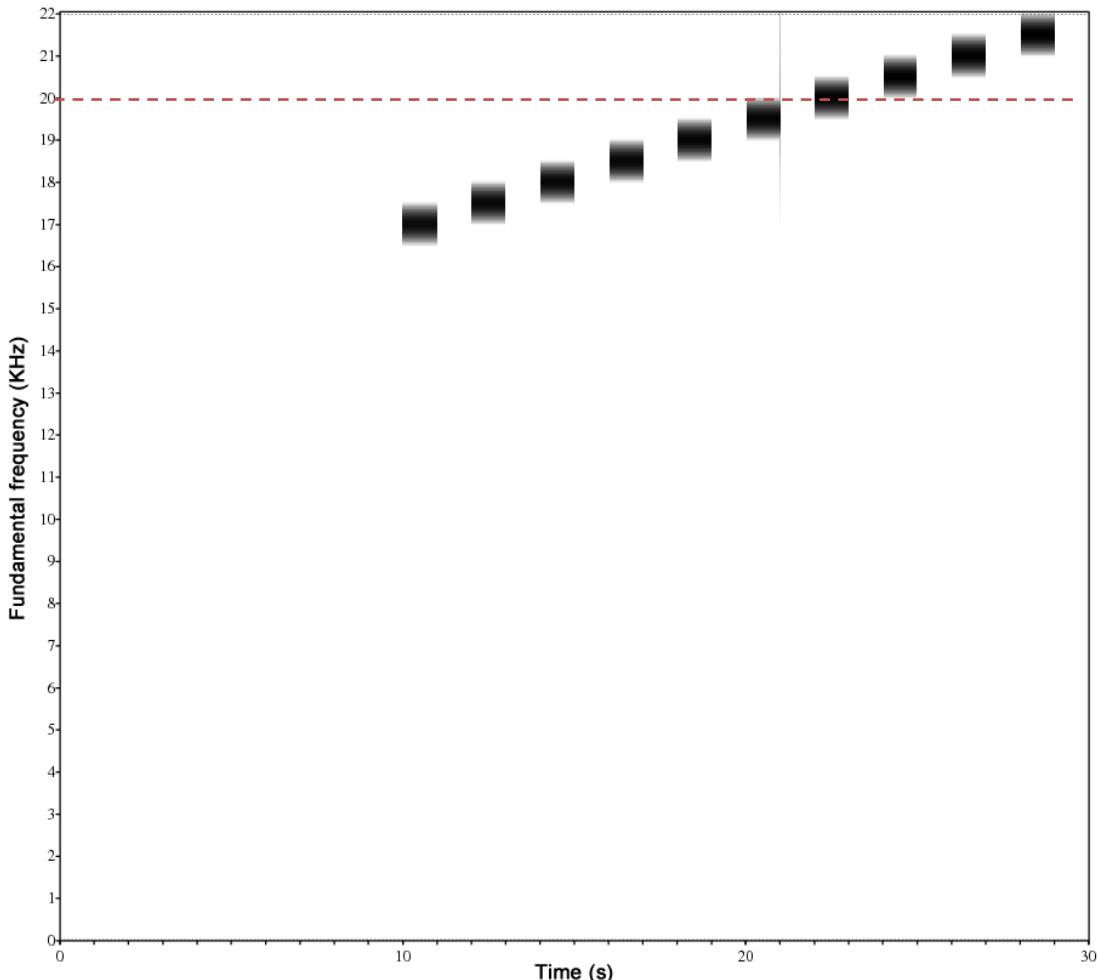
In order to eliminate any incidental/background noise, the ultrasound generation experiments were done in a soundproof recording booth. The recording was done using one Neumann U 87 Ai and Pro Tools software. The setup can be found in Figure 12.



**Figure 12: The ultrasound generation experiment setup.** The double bold square represents a sound proof recording booth. The slim single line represents the connection between the microphone and the Pro Tools system.

There are very few microphones that officially detect frequencies up to 22 kHz. A majority of professional microphones officially cover 20 Hz to 20 kHz, with cheaper models sometimes stopping at 17 kHz. This is only a precaution and microphones have been known to capture frequencies above the upper limit given in the specifications. Since microphone specifications cannot be relied on, it is necessary to confirm that the chosen microphone can detect signals in the entire range before each of the mobile phones can be tested. Neumann U 87 Ai microphone was successfully tested by playing

one of the sound files, described later in this section, through Beyerdynamic DT 150 earphones at high volume. The specifications for these earphones state they can produce frequencies up to 30 kHz.



**Figure 13: A spectrogram of the file played back by all tested smartphones.** X axis depict time and Y axis depict Frequency. Chromatic value shows energy (darker colour represents more energy). The dotted red line shows the upper human hearing range

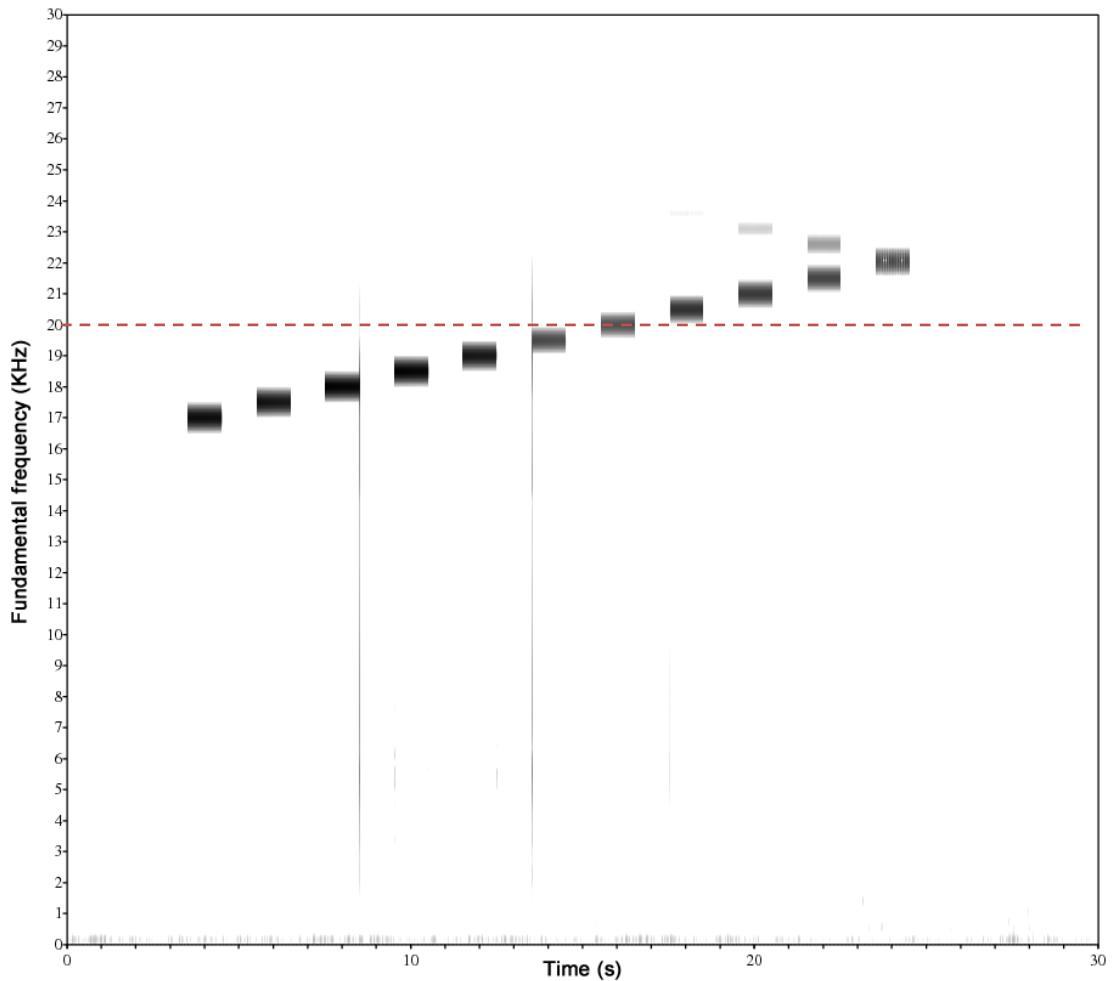
Initially one 44.1 kHz “WAV” sound file was generated using WaveLab software. This file starts with 10 seconds of silence in order to allow enough time to place the phone in front of the microphone, close the door of the recording booth and start recording. These ten seconds are followed by 11 one second long signals ranging from 17 to 22 kHz with a half kHz step. There is a gap of one second between each signal. A spectrogram of

this file can be seen on Figure 13, where darker colour represents more energy and the dotted red line shows the upper human hearing range.

During the early stages of the experiment it was observed that mobile phones generate a lot of noise in the lower frequencies when playing some or all of the given signals at maximum volume. This effect fades or disappears differently on different devices when volume is decreased. To reflect on that the testing procedure was modified. First of all four more modifications of the sound file were generated where volume is decreased by 20, 40, 60 and 80 percent. Secondly each of the five files was played at maximum volume on the device as well as one and two steps lower from maximum. This resulted in 15 separate recordings per device or 60 altogether. A spectrogram was generated for each of the 60 recordings using Praat software for further analysis. All spectrograms can be found in Appendix 1.

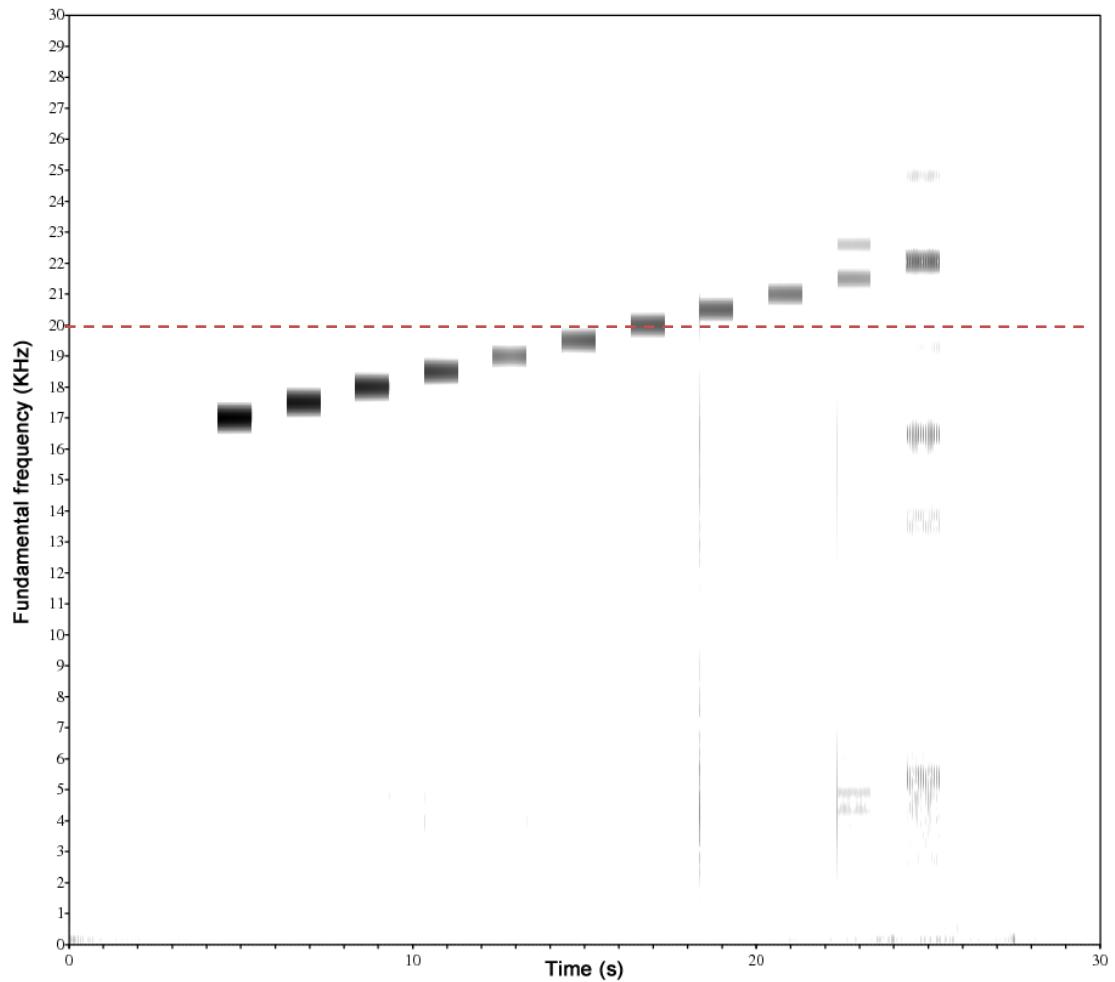
**Discussion.** Based on the spectrograms generated during the experiment the following observations were made:

1. All tested devices are able to generate all of the given frequencies under the condition that the volume is not too high. In other words there was always energy in the part of the histogram corresponding to the signal. Also for every device it is possible to find a volume setting at which the spectrogram looks almost the same as the spectrogram of the original file. For example with G1 the settings will be file volume 80%, device volume maximum - 2 .(Figure 14)



**Figure 14: Spectrogram for HTC G1 at file volume 80%, device volume maximum – 2.**

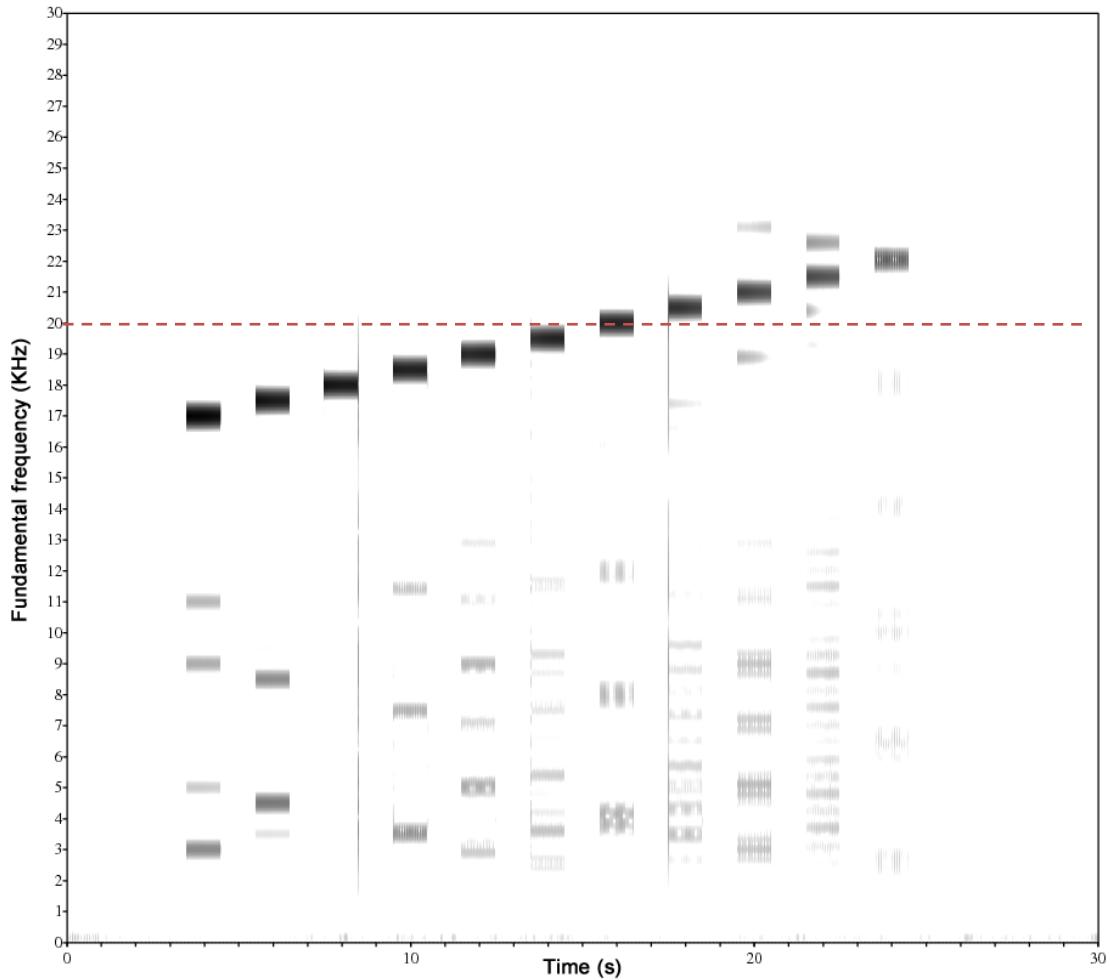
2. If the volume is set too high, mobile phones will generate a lot of noise in a wide range of frequencies in the audible range when trying to generate one of the signals. For iPhone this happens only with 21.5 and 22 kHz, but for Hero and Navigator this happens at all tested frequencies. (See Figures 15 and 16.) Only HTC G1 appeared to be almost completely immune to this problem. As the volume is decreased, this problem fades, and at some point disappears. For example with HTC Hero this happens at around 80% file volume at maximum device volume. With iPhone noise at 21.5 and 22 kHz disappears completely around 20% file volume and device volume maximum - 2.



**Figure 15: Spectrogram for iPhone at file volume 60%, device volume maximum.**

3. Volume settings of the device have a major impact on the appearance of noise.

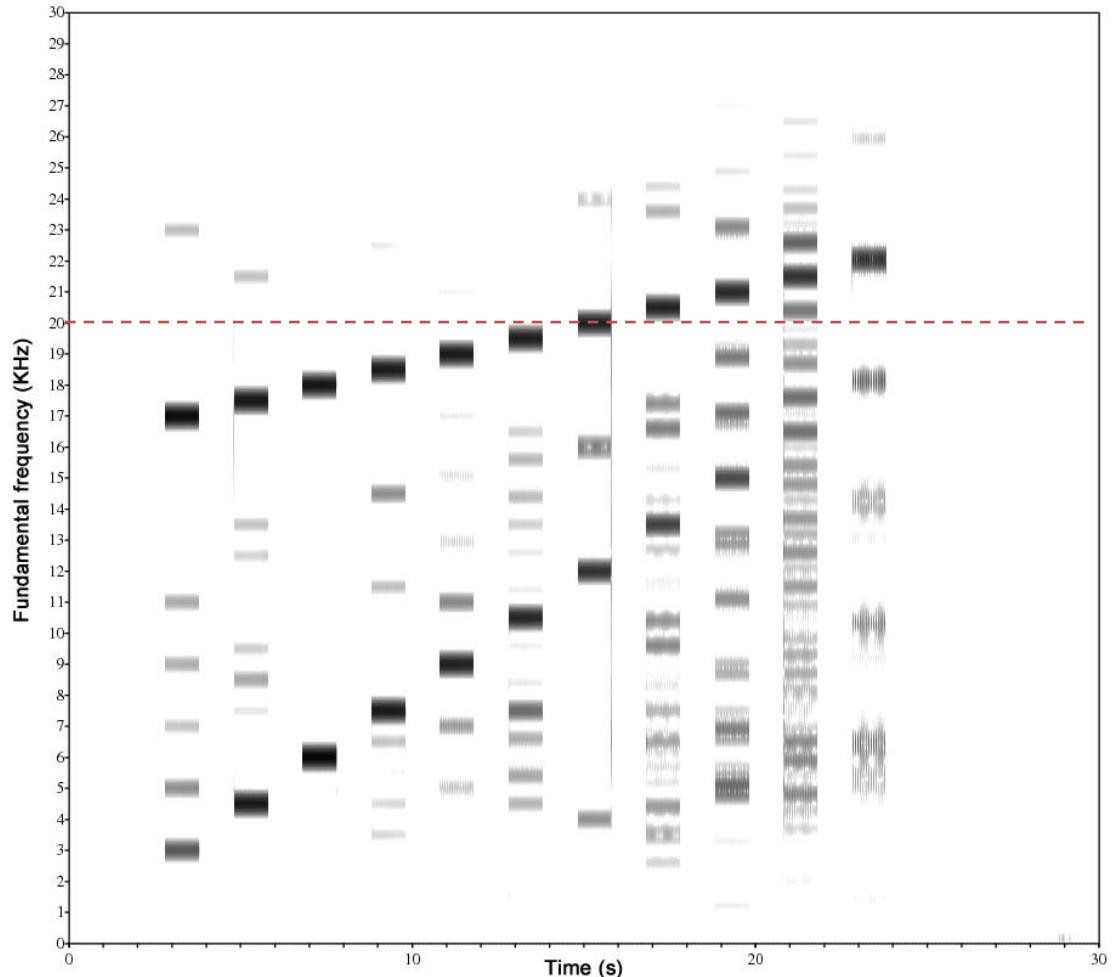
This was particularly vivid with Nokia Navigator, where it was impossible to avoid noise even with 20% file volume. Noise almost completely disappeared when the device was set to maximum - 2 even with 100% file volume. With other devices it was only observed that noise can be almost completely eliminated by setting the device volume only one or two steps lower than maximum. Reducing volume in the file seemed to have less impact. (See Figure 17 and 18 for comparison.)



**Figure 16: Spectrogram for HTC Hero at file volume 100%, device volume maximum**

4. In a majority of recordings there can be observed a particular pattern of phantom signals which are a few kHz higher than the real signal. Sometimes they are almost as powerful the real signal, but very often are hardly visible. A very vivid example can be seen on Figure 18, but for other phones the effect is close to Figure 14. This is probably caused by either resonance in speaker diaphragm or operational errors in Digital Signal Processing (DSP) hardware. This trend may impact scalability of the positioning system. For example as can be seen on Figure 14, the system wouldn't be able to tell whether the original signal was 21.5 or 22.5 kHz. If two different devices used these different frequencies to uniquely identify themselves,

the system would fail to tell whether the two signals are an original and a phantom or two simultaneous signals from the two devices.

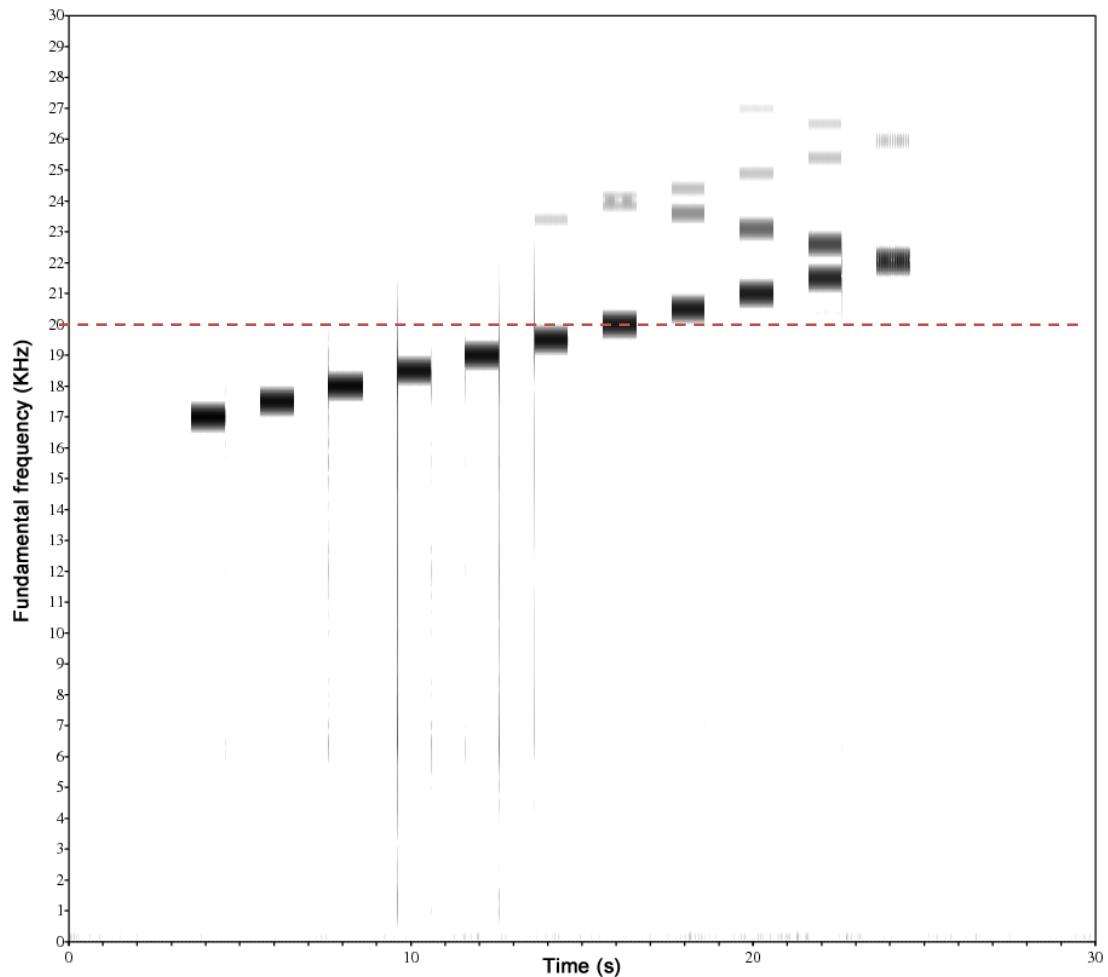


**Figure 17: Spectrogram for Nokia Navigator at file volume 20%, device volume maximum.**

*There is a lot of noise even despite very low volume of the signal in the file.*

With the exception of very high volume settings, all tested mobile phones performed generation of 17-22 kHz signals very well. Some devices performed better than others. HTC G1 didn't generate almost any audible noise even at the highest settings. iPhone showed even less noise at the highest settings with the exception of 21.5 and 22 kHz signals. The other two phones generated a lot of noise at the highest volume settings. The problem with audible noise being generated along with ultrasound was easily avoided by reducing the volume settings on the device. Making the original signal

quieter seemed to have less effect or even no effect at all as on Nokia 6210 Navigator. On most devices 20-22 kHz signals were accompanied by noise in the upper frequencies as on Figure 18. Reducing signal volume didn't have almost any effect on them. Although this noise is unavoidable it will not have any impact on usability being inaudible, but should be taken into consideration when scaling up the system to accommodate more devices. From our observations we can conclude that the cause of the noise in the upper frequencies is different from the cause of noise in lower frequencies.

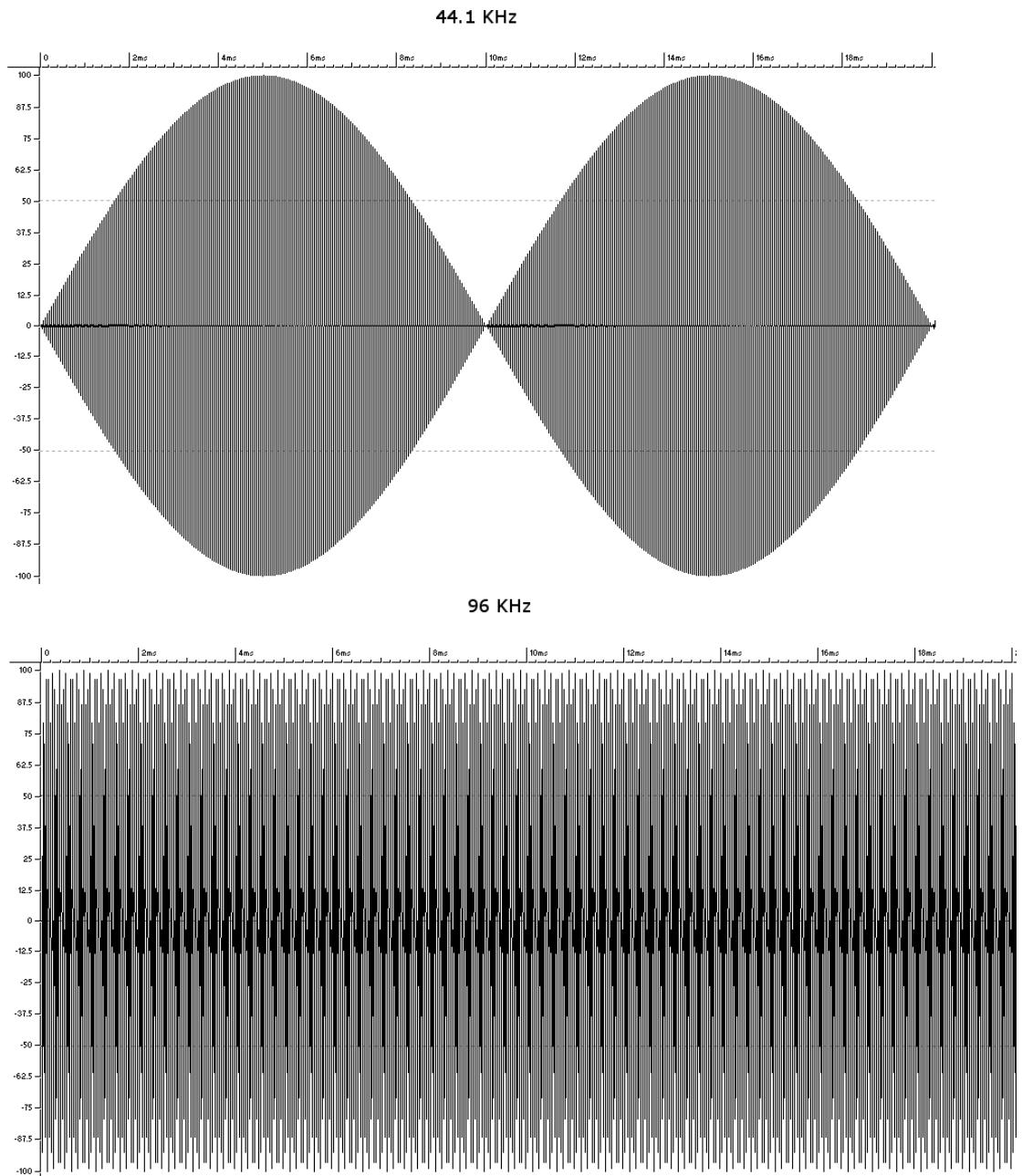


**Figure 18: Spectrogram for Nokia Navigator at file volume 100%, device volume maximum - 2.**  
*Audible noise abruptly disappears at maximum - 2 settings even though file volume is high.*

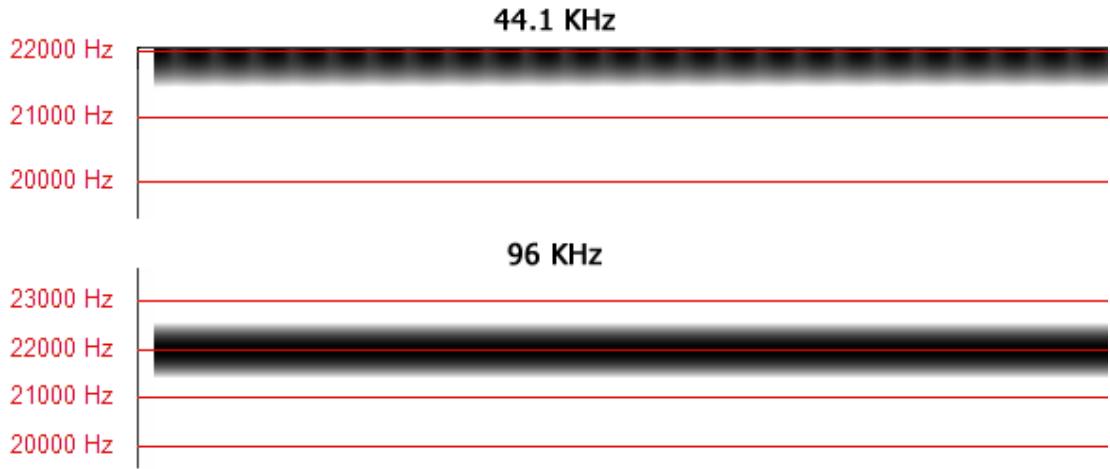
### 3.2. Signal Design

This section addresses the second research question (RQ2): What are the desirable characteristics of the emitted signal? In order to answer this question, ultrasound samples of various shapes, lengths and frequencies were created and tested for suitability.

**Frequency.** There is an upper and lower limit imposed on the frequency the ultrasonic signal may occupy. The lower limit is 20 kHz – the upper limit of human hearing range. Because young people can to some extent sense frequencies above this limit at close range, it is desirable to use higher frequencies rather than frequencies just above the threshold. The upper limit is less arbitrary as it is imposed by technical limitations of mobile phones that don't support sample rates above 44.1 kHz. According to the sampling theorem the sampling frequency must be twice the maximum frequency, which means that mobile phones can't produce frequencies above 22 kHz. In reality the boundary for seamless playback is lower. Due to the way high frequencies are stored in 44.1 kHz files a simple uniform high-frequency signal will have drops in energy down to zero with a certain periodicity. This periodicity depends on the chosen frequency and quickly drops as signal frequency approaches the upper boundary. While at 21.5 kHz drops occur every millisecond and don't distort the intended shape of the signal too much, at 22 kHz they occur every 10 ms and distortion is significant. The difference between the same 22 kHz wave stored at 44.1 and 96 kHz sampling rate is shown on Figure 19 and 20. Also a lower frequency will help make the system compatible with running microphones at 44.1 kHz as using frequency very close to the upper limit may cause aliasing.



**Figure 19: Waveforms of the same 22kHz signal saved at 44.1 and 96 kHz sampling rate. Although both waveforms represent the same frequency, the top one looks different due to aliasing.**



**Figure 20:** Spectrograms of the same 22kHz signal saved at 44.1 and 96 kHz sampling rate. The top graph shows a spectrogram of a 22 kHz wave stored at 44.1 kHz sampling rate. The bottom graph shows the same frequency at 96 kHz sampling rate. 44.1 kHz sampling rate has regular drops in energy that are not present at 96 kHz. The length of the signal shown on the diagram is 250 milliseconds.

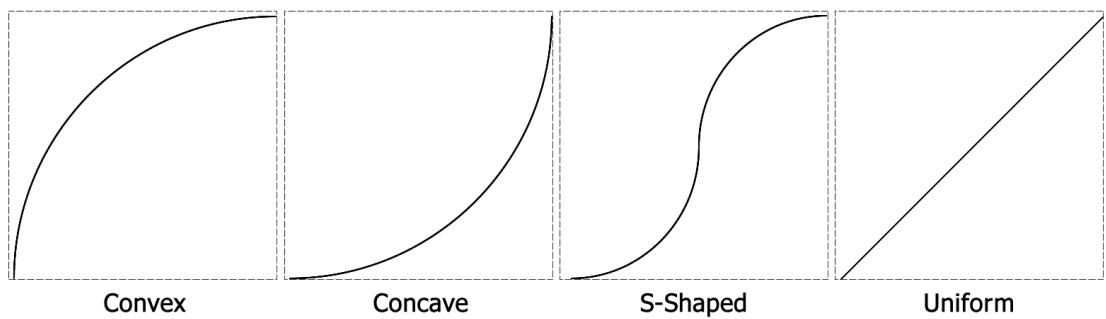
For a system that uses only one frequency it can be concluded that any frequency between 20.5 and 21.5 is a good choice. Three identically shaped signals; 20.5, 21 and 21.5 kHz, were tested on HTC G1 mobile phone and no noticeable increase in audibility was observed in any of the three samples. For systems that require several frequencies, a balance will have to be found between tradeoffs described above and having to deal with closely aligned frequencies. For the 20-22 kHz band a step of 300 Hz would be enough to reduce the imprint of the signal down to 1% of its original energy in neighbouring bands after applying a band pass filter. It should be noted however that the shape and size of this imprint is similar to a very badly attenuated signal. In other words a very strong 21.3 kHz signal after a check for 21.6 kHz signal may be confused with a very weak 21.6 kHz signal. If this were to happen, the result would a false positive.

**Shape and length.** In order to maximise effective distance, the portion of signal used for detection must have the highest possible volume. However, just this one chirp without a gradual increase (attack) and decrease (release) in volume produces an unwanted audible clicking noise. A number of different envelopes varying in shape and

length of both fade in and fade out slopes were tested. The following requirements were addressed:

- The signal must be as difficult to detect with the human ear as possible.
- The slopes must not obstruct detection of the point of reference, which they envelope.
- The entire signal should be as short as possible, to reduce the time it takes to playback the signal, process it, and overall reduce position calculation lag.

Convex and S-shaped slopes were not tested because gradual slope close to the top obstructs precise identification of the peak. Uniform slopes failed to eliminate audible clicks. But concave slopes worked very well, and the stronger the curve the softer the signal sounded. Figure 21 shows what convex, concave, s-shaped and uniform slopes look like. In the final signal we used a concave slope with 90% offset which didn't just perform well in terms of inaudibility but also helped detection as it made the reference point more prominent.



**Figure 21: Common slope shapes.**

The length of the attack slope was also tested. While 5 and 10 millisecond slopes presented sharp loud clicks 15 ms slope behaved unpredictably and produced a loud

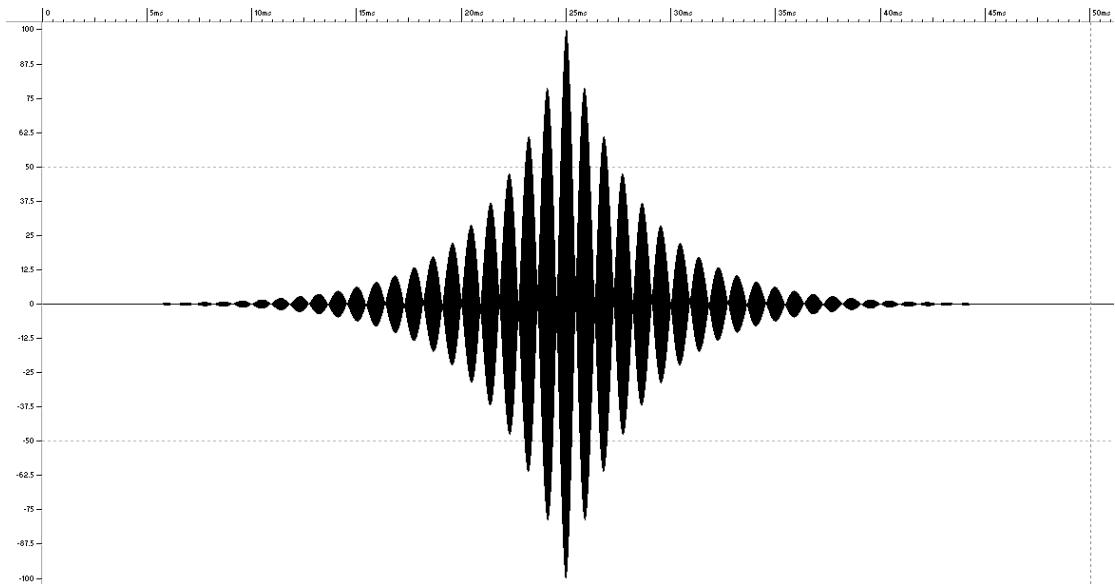
audible click roughly one out of every five times. 25 ms slope managed to consistently eliminate sharp audible noises and was therefore kept for the final design.

As for the release slope, it was confirmed that it is just as necessary to eliminate distortions. Attempts to make it shorter than the attack slope were unproductive so in the final design it was also set to 25 ms. This makes the length of the entire signal 50 ms which interestingly matches recommendations made by Peng et. al. even though they had different requirements and worked with audible sound (Peng 2007). In addition, ultrasonic signal being symmetric makes finding the reference point at high levels of attenuation easier. Because the peak quickly becomes less prominent as the signal attenuates, we will be sometimes presented with some energy spread over a period of time. With a symmetric signal, it can be safely assumed that the reference point is roughly in the centre. This of course can be hindered by energy arriving via multipath, but still not as bad as if the reference point was not in the centre. In this case it is impossible to tell to what extent a certain part of a highly attenuated signal is energy from the peak offset in time and to what extent it is remainder of the envelope. Optimal single-frequency waveform can be seen on Figure 22.

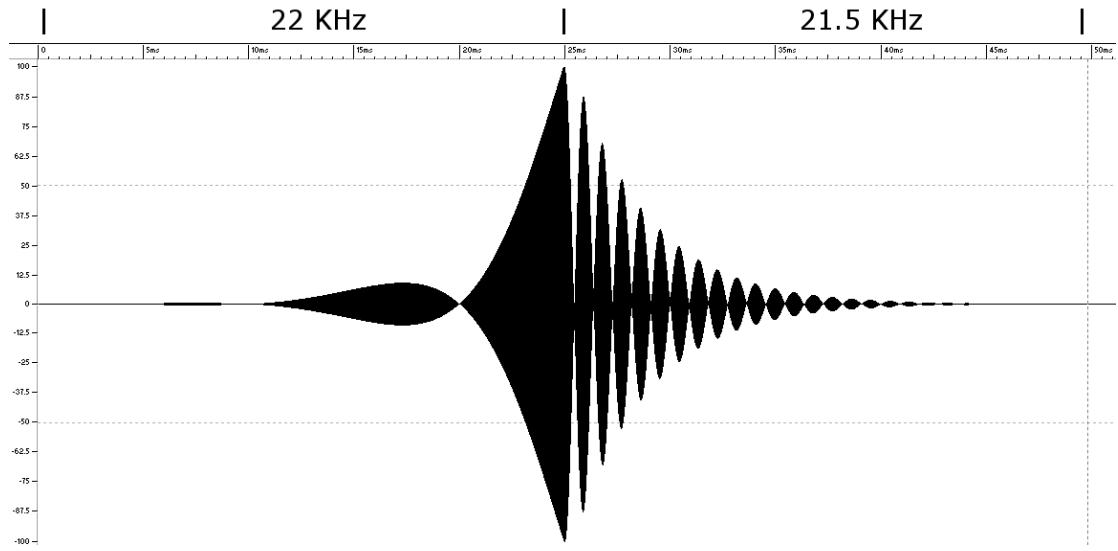
**Frequency alternation.** By making the envelope completely or partially out of a different frequency it is possible to reduce the envelope's footprint in the filtered frequency and make the signal itself more pronounced. Theoretically the right side of the envelope is not important, because when recorded it will merge with sound that arrived via multipath. By replacing the first half of the signal with a different frequency (as seen on Figure 23) we make it harder to confuse the beginning of the signal with the envelope. An alternative solution is to replace the entire envelope with a different frequency as seen on Figure 24. We found that this second solution is easier to work with. One limitation is that the two frequencies should be joined at the top of the

amplitude. Joining at the bottom of the amplitude will cause an audible noise. It is possible to keep the signal part fairly small thanks to the use of high frequencies.

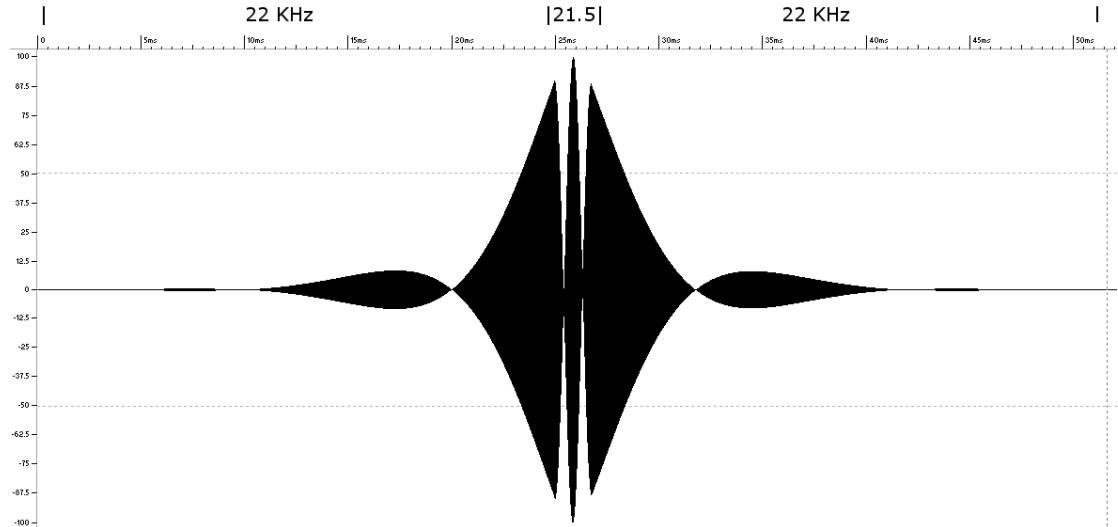
**Volume.** Thanks to other properties of the signal, it was possible to use the signal at maximum volume without distortions, thus maximizing effective distance as a result.



**Figure 22:** Waveform of the optimal single-frequency signal. The X-axis depicts time in milliseconds and the Y-axis is energy level in percentage. The signal actually occupies all of 50 ms, but the energy in front and at the end of the graph is too weak to be seen at this scale.



**Figure 23:** Waveform of the signal where the first half is composed using a different frequency. The X-axis depicts time in milliseconds and the Y-axis is energy level in percentage. Used frequency is shown at the top.



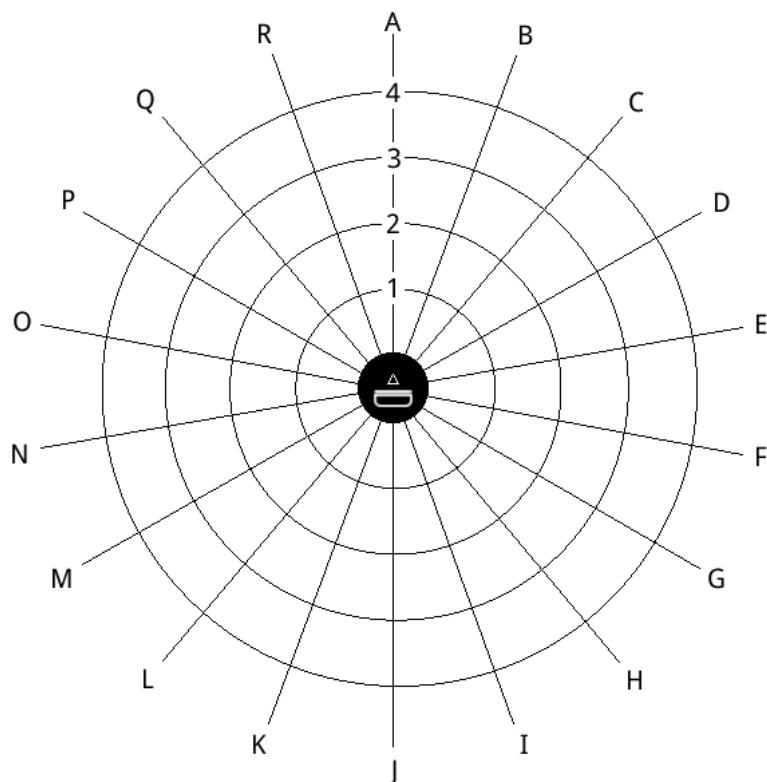
**Figure 24:** Waveform of the signal where the entire envelope is composed using a different frequency. The X-axis depicts time in milliseconds and the Y-axis is energy level in percentage. Used frequency is shown at the top.

### 3.3. Range

This section addresses the third research question (RQ3): What is the maximum distance at which an ultrasound signal emitted by a mobile phone can be reliably detected with a microphone? While sound propagates with the same speed in all directions, the same cannot be said about intensity. Ultrasound is highly directional, meaning that sound intensity in front of the speaker is much higher than to the sides and behind the speaker. This means maximum distance at which the signal can be detected is significantly longer if the microphone is positioned in the direction the speaker faces compared to any other direction. Since the user is free to move and rotate the phone at random, maximum functional distance will be the longest distance a signal can be detected from any direction. It should be sufficient to only test scenarios where a microphone is oriented to face the speaker, because in the positioning setup microphones will be placed in corners facing the centre of the room.

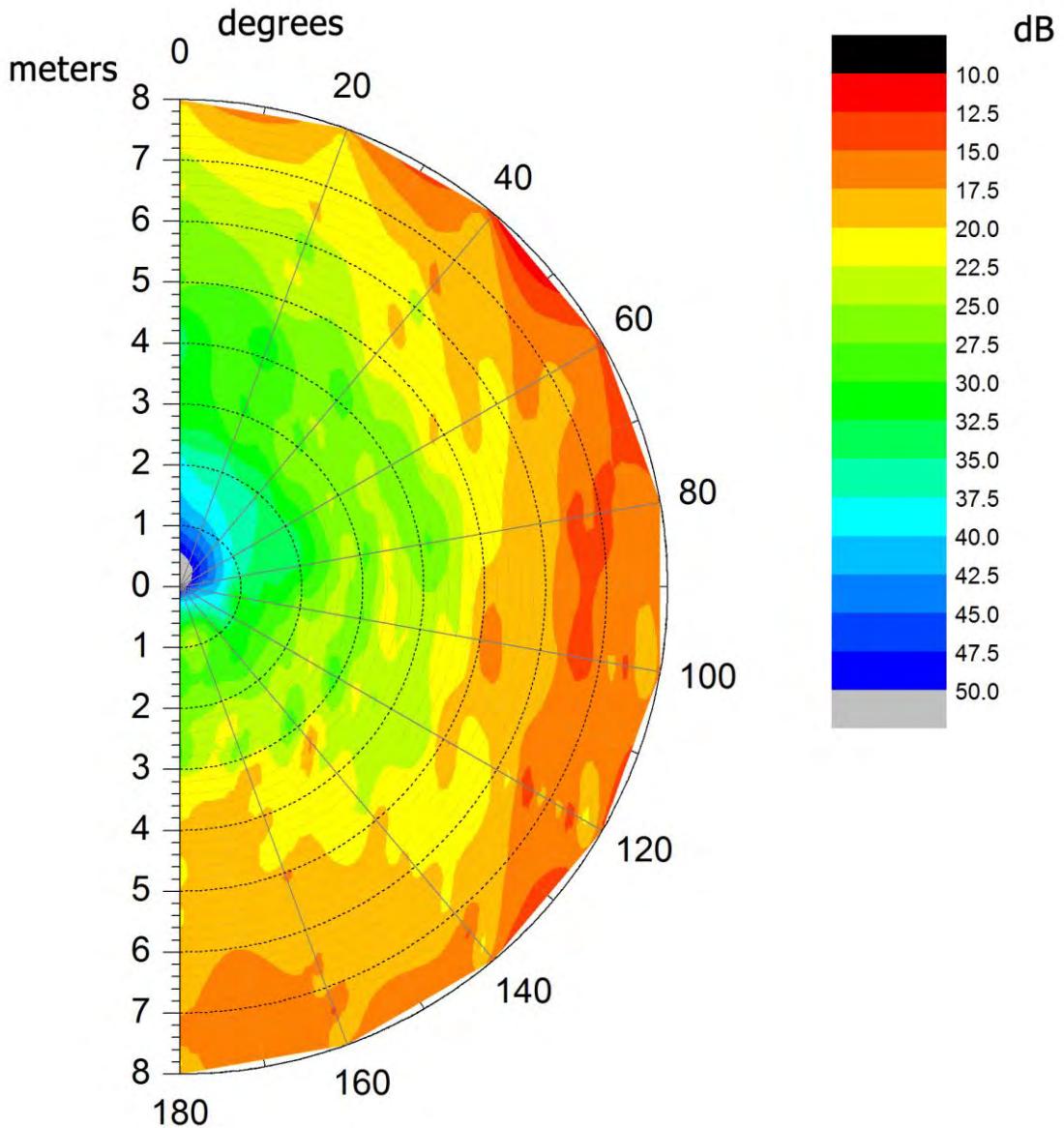
**Experiment Method.** For this range experiment the HTC G1 phone was placed in the centre of a large room. The area around the phone was divided into eighteen 20 degree segments, as shown in Figure 25. Each line is identified by a letter. The area is further divided by equally spaced concentric circles, identified by numbers. The spacing between circles is 20 centimetres. There are 40 of these circles making the radius of the area covered by the experiment to be 8 meters. Because the area directly in front of the speaker was expected to show significantly higher energy readings, additional recordings were made just for line A up to 20 meters with a 1 meter step. A combination of a letter and a number were used to identify each unique location – an intersection of a line and a circle. Signal intensity at a given direction and distance was measured by placing a microphone at a corresponding location. The microphones were all set at the same height as the mobile phone using a meter tape. The precise distance

and angle were verified using a laser measuring device. Number of the recording was written down next to the letter and number combination so that the location of the microphone could be recovered. Because signal propagation has rotational symmetry around the axis pointing in the direction the speaker is facing recordings were done only for sectors A-J. Altogether 400 recordings were made plus additional 12 recordings for line A as mentioned above. All the signals were recorded at 24/192 kHz using a TASCAM HD-P2 portable recording system.



**Figure 25: The division of space around the phone into sectors.** Letters are used to define direction and numbers to define distance to the phone. Combination of a letter and a number uniquely identifies a microphone's position.

A pass Hann band filter (Dixon 1983) was applied to each of the recordings in order to isolate the 22 kHz frequency component. After that the highest intensity value was extracted for each of the files. These values were used to generate a polar contour plot shown in Figure 26. All values can be found in Appendix 2.



**Figure 26: Polar contour plot for ultrasound energy propagation.** The semicircle represents a top-down view on half of the area around the phone. Chromatic value represents energy level in dB, for which corresponding numeric values can be seen in the legend.

**Discussion.** Based on the contour plot the following observations were made:

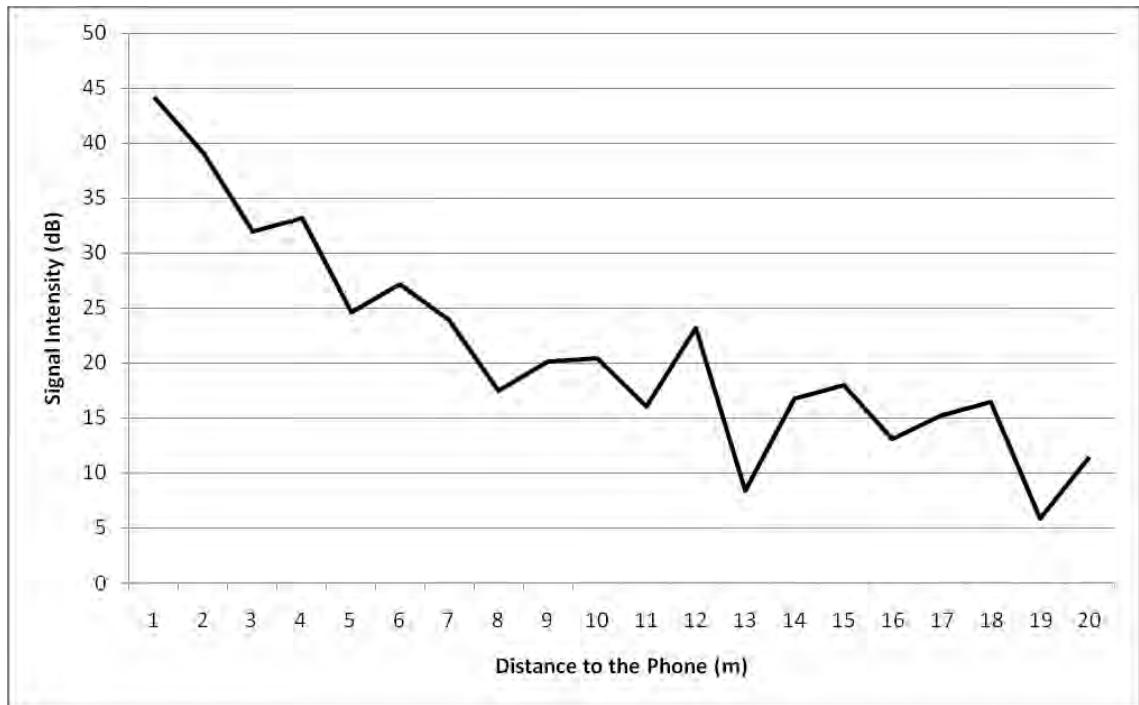
1. Sound intensity propagation for levels above 40 dB follows the shape of a cardioid (Figure 8, left). This is similar to the shape characteristic for 16 kHz on Figure 24. Depending on definition the shape can also be described as supercardioid. Very often supercardioid is depicted as on Figure 8 (right), but sometimes it means a

more elongated version of cardioids, which is very close to what we see on the contour plot.

2. Distribution of energy below 40 dB is less predictable. Although it can be observed that it resembles a subcardioid (Figure 8, bottom), there is a large element of randomness. There may be two factors contributing to that. First, reflected resonance within the phone due to various metal and plastic parts vibrating and interacting with each other. And secondly, finite microphone sensitivity.
3. In the entire experiment only one value below 10 dB was registered. This is well above the 21.5 kHz component of background noise which is around 1 dB. This means that the signal can be detected from any direction within 8 meter radius. However there is no guarantee that the maximum value belongs to a signal that arrived by direct path and not via a longer path.

Additional information collected for sector A up to 20 meters was used to generate a graph that depicts relationship between distance and signal strength (see Figure 27). It can be seen that even with a speaker and a microphone pointing at each other, signal strength can't be directly used to estimate distance.

Overall the experiment demonstrated that an ultrasound signal produced by a mobile phone can be detected at a distance at least up to 8 meters regardless of phone's orientation. This suggests that for a room with a diagonal of 8 meters it should be sufficient to place a microphone in each corner. This way a phone placed anywhere in the room will always within 8 meter range from all four microphones.



**Figure 27: Relationship between signal strength and distance for conditions where phone speaker and microphone point at each other.**

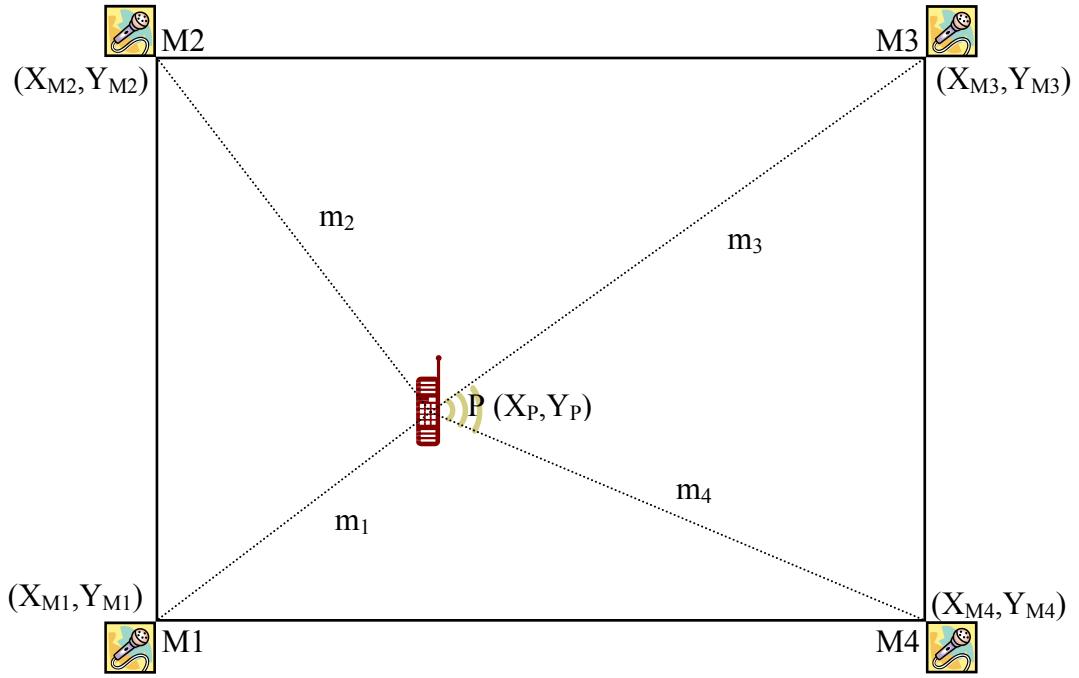
## 4. ASYNCHRONOUS TRILATERATION

This chapter addresses the fourth research question (RQ4): Can ultrasound positioning be done asynchronously? Using trilateration it is possible to calculate one's position based on the distance to several other (control) points with known positions (Bossler 2002; Ghilani 2006). To find one's position in 2 dimensions the number of required known points is 3; for position in 3 dimensions the number of known points is 4. Given that the speed of sound propagation is constant under the same temperature and humidity conditions, the time it takes a signal to travel between the phone speaker to each known microphone control point (in our case) can be directly converted into distance between the phone and microphones. This is the TOA (Time of Arrival) approach. In general, the main problem with this approach is that both the time the signal was sent and the time it was received are required in order to get the time-of-flight.

In our scenario of quickly and accurately locating a mobile phone indoors, TOA requires that times from two separate systems with two separate clocks will have to be synchronised - a major source of error. As such it is desirable to compare only the time of arrival at each of the microphones and ignore completely the time the signal was originally sent from the phone, making Lok8 a TDOA (Time Difference of Arrival) approach.

### 4.1. Least Squares Method for TOA Trilateration

To help illustrate the problem we first introduce the concept of Least Squares for solving TOA trilateration where all distances between the phone and microphones are known. (See Figure 28). Adapted from (Ghilani 2006).



**Figure 28: Time of Arrival.** Control points  $M1, M2, M3$  and  $M4$  are known microphone positions. Point  $P$  is the unknown mobile phone's position in a room, coordinates of which we are trying to find. Lines  $m1, m2, m3$  and  $m4$  are known distances between the phone and each microphone.

From Pythagoras we derive the following *mathematical model* to describe the ultrasonic relationships between phone  $P$  and microphones  $M1, M2, M3, M4$ :

$$m_1^2 = (X_P - X_{M1})^2 + (Y_P - Y_{M1})^2 \quad \text{or} \quad m_1 = \sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2}$$

$$m_2^2 = (X_P - X_{M2})^2 + (Y_P - Y_{M2})^2 \quad \text{or} \quad m_2 = \sqrt{(X_P - X_{M2})^2 + (Y_P - Y_{M2})^2}$$

$$m_3^2 = (X_P - X_{M3})^2 + (Y_P - Y_{M3})^2 \quad \text{or} \quad m_3 = \sqrt{(X_P - X_{M3})^2 + (Y_P - Y_{M3})^2}$$

$$m_4^2 = (X_P - X_{M4})^2 + (Y_P - Y_{M4})^2 \quad \text{or} \quad m_4 = \sqrt{(X_P - X_{M4})^2 + (Y_P - Y_{M4})^2}$$

Re-write above four mathematical model equations as *observation equations* by adding a residual  $v_m$  (random error) to each measurement:

$$F: m_1 + v_{m1} = \sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2}$$

$$G: m_2 + v_{m2} = \sqrt{(X_P - X_{M2})^2 + (Y_P - Y_{M2})^2}$$

$$H: m_3 + v_{m3} = \sqrt{(X_P - X_{M3})^2 + (Y_P - Y_{M3})^2}$$

$$I: m_4 + v_{m4} = \sqrt{(X_P - X_{M4})^2 + (Y_P - Y_{M4})^2}$$

Because number of measurements ( $m = 4$ ) is greater than number of unknowns ( $n = 2$ ), use Least Squares to determine the MPV of the unknowns ( $X_P, Y_P$ ). Since the observation equations are non-linear in the unknowns ( $X_P, Y_P$ ), a first-order *Taylor Series* is needed to approximate a set of linear observation equations before taking partial derivatives.

Considering function F above (describing ultrasonic relationship between M1 and P):

$$F: m_1 + v_{m1} = \sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2}$$

This non-linear function can be written as:

$$F(X_P, Y_P) = m_1 + v_{m1}$$

Where

$$F(X_P, Y_P) = \sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2}$$

The above function is *linearized* using a first-order Taylor Series approximation:

$$F(X_P, Y_P) = F(X_{P_o}, Y_{P_o}) + \left( \frac{\partial F}{\partial X_P} \right)_o dX_P + \left( \frac{\partial F}{\partial Y_P} \right)_o dY_P$$

Where

- $X_{P_o}$  and  $Y_{P_o}$  are initial estimates of smartphone position in the room calculated by taking average of all known microphone positions.
- $F(X_{P_o}, Y_{P_o})$  is the non-linear function evaluated with these estimates.
- $dX_p$  and  $dY_p$  are corrections to the initial estimates such that  $X_p = X_{P_o} + dX_p$  and  $Y_p = Y_{P_o} + dY_p$

The partial derivatives  $\left( \frac{\partial F}{\partial X_P} \right)$  and  $\left( \frac{\partial F}{\partial Y_P} \right)$  are found by first re-writing function F:

$$F: \quad F(X_P, Y_P) = \sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2}$$

and then take partial derivative with respect to  $X_P$ :

$$\begin{aligned} \frac{\partial F}{\partial X_P} &= \frac{1}{2} \sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2} \bullet 2(X_P - X_{M1}) \\ &= \frac{(X_P - X_{M1})}{\sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2}} \\ &= \frac{(X_P - X_{M1})}{m_1} \end{aligned}$$

and then with respect to  $Y_P$ :

$$\frac{\partial F}{\partial Y_P} = \frac{1}{2} \sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2} \bullet 2(Y_P - Y_{M1})$$

$$= \frac{(Y_P - Y_{M1})}{\sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2}}$$

$$= \frac{(Y_P - Y_{M1})}{m_1}$$

Therefore:

$$F(X_P, Y_P) = F(X_{P_o}, Y_{P_o}) + \left( \frac{(X_P - X_{M1})}{m_1} \right)_o dX_P + \left( \frac{(Y_P - Y_{M1})}{m_1} \right)_o dY_P$$

So the linearized observation equation for  $m_1$ , describing the ultrasonic relationship between microphone M1 and phone P becomes:

$$\left( \frac{(X_P - X_{M1})}{m_1} \right)_o dX_P + \left( \frac{(Y_P - Y_{M1})}{m_1} \right)_o dY_P = (m_1 - m_{1_o}) + v_{m1}$$

where  $m_{1_o}$  is the initial estimate of  $m_1$  calculated using  $X_{P_o}, Y_{P_o}$ .

Likewise for function G (between M2 and P):

$$\left( \frac{(X_P - X_{M2})}{m_2} \right)_o dX_P + \left( \frac{(Y_P - Y_{M2})}{m_2} \right)_o dY_P = (m_2 - m_{2_o}) + v_{m2}$$

function H (between M3 and P):

$$\left( \frac{(X_P - X_{M3})}{m_3} \right)_o dX_P + \left( \frac{(Y_P - Y_{M3})}{m_3} \right)_o dY_P = (m_3 - m_{3_o}) + v_{m3}$$

and function I (between M4 and P):

$$\left( \frac{(X_P - X_{M4})}{m_4} \right)_o dX_P + \left( \frac{(Y_P - Y_{M4})}{m_4} \right)_o dY_P = (m_4 - m_{4o}) + v_{m4}$$

When using Matrix Methods for Least Squares, the observation equations are represented in matrix form as:

$${}_m A_n n X_1 = {}_m L_1 + {}_m V_1$$

Where in our case:

- $m = 4, n = 2$
- ${}_m A_n$  contains the coefficients of the unknowns  $(X_P, Y_P)$
- ${}_n X_1$  contains the corrections to be applied to the initial estimates for the unknowns  $(dX_P, dY_P)$
- ${}_m L_1$  contains the measurements  $(m_1, m_2, m_3, m_4)$
- ${}_m V_1$  contains the residuals (one for each measurement)

Solving for  $X$  gives the solution:

$$X = (A^T A)^{-1} A^T L \text{ where:}$$

$$A = \begin{bmatrix} \frac{(X_P - X_{M1})}{m_1} & \frac{(Y_P - Y_{M1})}{m_1} \\ \frac{(X_P - X_{M2})}{m_2} & \frac{(Y_P - Y_{M2})}{m_2} \\ \frac{(X_P - X_{M3})}{m_3} & \frac{(Y_P - Y_{M3})}{m_3} \\ \frac{(X_P - X_{M4})}{m_4} & \frac{(Y_P - Y_{M4})}{m_4} \end{bmatrix} \quad (4-1)$$

$$X = \begin{bmatrix} dX_P \\ dY_P \end{bmatrix} \quad (4-2)$$

$$L = \begin{bmatrix} m_1 - m_{l_0} \\ m_2 - m_{l_0} \\ m_3 - m_{l_0} \\ m_4 - m_{l_0} \end{bmatrix} \quad (4-3)$$

$$V = \begin{bmatrix} v_{m1} \\ v_{m2} \\ v_{m3} \\ v_{m4} \end{bmatrix} \quad (4-4)$$

Matrix  $X$  contains the corrections to be applied to the original estimates for  $(X_P, Y_P)$ .

These new  $(X_P, Y_P)$  coordinates are then used to recalculate updated distances for  $(m_{l_0}, m_{l_0}, m_{l_0}, m_{l_0})$ . The process is repeated until coordinates of  $(X_P, Y_P)$  don't change significantly (e.g. in the 3<sup>rd</sup> decimal place for mm precision).

After a solution has been reached, the residuals  $V$  for each measurement and Standard Deviation of *unit weight*  $\sigma_o$  for the overall least squares adjustment can be calculated with:

$$V = AX - L \quad \text{and} \quad \sigma_o = \pm \sqrt{\frac{(V^T V)}{r}}$$

Where *degrees of freedom*  $r = m-n$  and the Standard Deviation of each adjusted unknown is then given by:

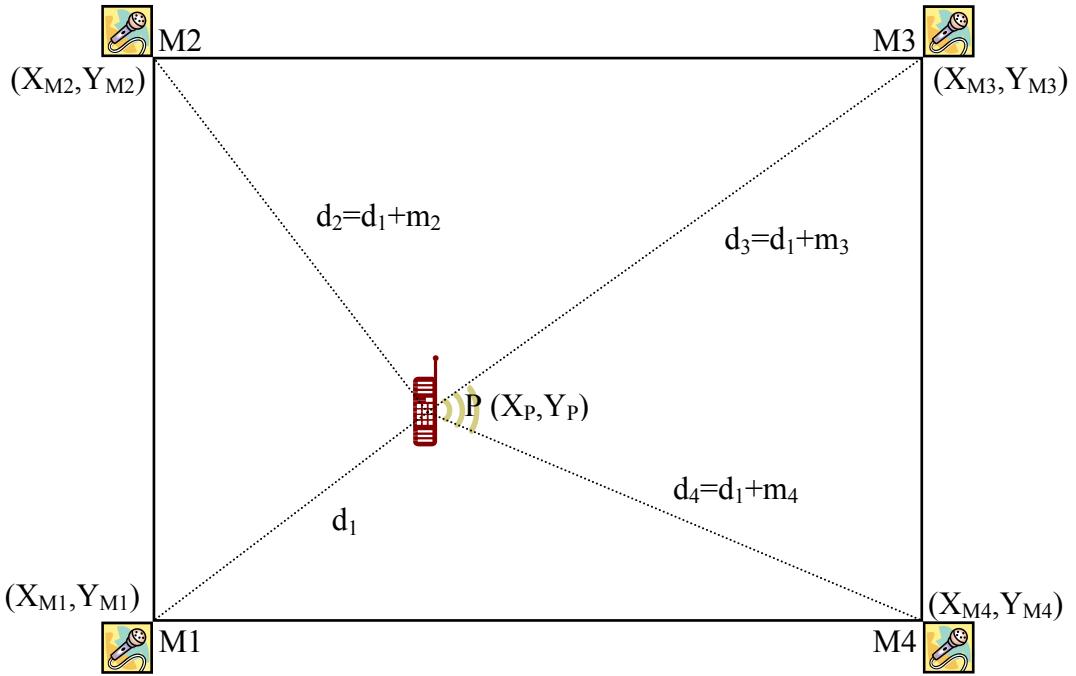
$$\sigma_{x_i} = \pm \sigma_o \sqrt(Q_{x_i x_i})$$

In our case  $\sigma_{X_1}$  is the Standard Deviation for  $X_P$ , and  $\sigma_{X_2}$  is the Standard Deviation for  $Y_P$ . These standard deviations imply that there is a 68% probability that the adjusted values for  $X_P$  and  $Y_P$  are within  $\pm \sigma$  of this amount.

$(A^T A)^{-1}$  is called the *variance-covariance* matrix or  $(Q_{XX})$  matrix and  $(Q_{x_i x_i})$  is the *variance* of unknown  $i$ , or the element in the  $i^{th}$  row and  $i^{th}$  column of the  $(A^T A)^{-1}$  matrix.

## 4.2. Least Squares Method for TDOA Trilateration

The above standard TOA approach has been extended to allow for TDOA trilateration where we don't know any of the distances between the phone and microphones, but instead only the differences in time that the phone signal was received at each microphone location. The combination of this algorithm applied to indoor positioning on COTS mobile phones using ultrasound is not found in the literature, which makes our approach a contribution to the state-of-the-art in this research field.



**Figure 29: Time Difference of Arrival.** Control points  $M1$ ,  $M2$ ,  $M3$  and  $M4$  are known microphone positions. Point  $P$  is the unknown mobile phone's position, coordinates of which we are trying to find. Lines  $d1$ ,  $d2$ ,  $d3$  and  $d4$  are unknown distances between the phone and each microphone. However, what are known are the differences between the three time measurements, where time is converted to distance to produce:  $m2$ ,  $m3$  and  $m4$ .

The TDOA problem is illustrated in Figure 29 and the detailed solution follows. Java source code for TDOA trilateration can be found in Appendix 3.

From Pythagoras we derive the following *mathematical model* to describe the ultrasonic relationships between phone P and microphones M1, M2, M3, M4:

$$d_1^2 = (X_P - X_{M1})^2 + (Y_P - Y_{M1})^2 \quad \text{or} \quad d_1 = \sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2}$$

$$d_2^2 = (X_P - X_{M2})^2 + (Y_P - Y_{M2})^2 \quad \text{or} \quad d_2 = \sqrt{(X_P - X_{M2})^2 + (Y_P - Y_{M2})^2}$$

$$d_3^2 = (X_P - X_{M3})^2 + (Y_P - Y_{M3})^2 \quad \text{or} \quad d_3 = \sqrt{(X_P - X_{M3})^2 + (Y_P - Y_{M3})^2}$$

$$d_4^2 = (X_P - X_{M4})^2 + (Y_P - Y_{M4})^2 \quad \text{or} \quad d_4 = \sqrt{(X_P - X_{M4})^2 + (Y_P - Y_{M4})^2}$$

However, we can re-write  $d_2$ ,  $d_3$ ,  $d_4$  in terms of  $d_1$ :

$$d_2 = d_1 + m_2$$

$$d_3 = d_1 + m_3$$

$$d_4 = d_1 + m_4$$

And then substitute above  $d_1$  expressions back into the mathematical model:

$$d_1 + m_2 = \sqrt{(X_P - X_{M2})^2 + (Y_P - Y_{M2})^2} \text{ or } m_2 = \sqrt{(X_P - X_{M2})^2 + (Y_P - Y_{M2})^2} - d_1$$

$$d_1 + m_3 = \sqrt{(X_P - X_{M3})^2 + (Y_P - Y_{M3})^2} \text{ or } m_3 = \sqrt{(X_P - X_{M3})^2 + (Y_P - Y_{M3})^2} - d_1$$

$$d_1 + m_4 = \sqrt{(X_P - X_{M4})^2 + (Y_P - Y_{M4})^2} \text{ or } m_4 = \sqrt{(X_P - X_{M4})^2 + (Y_P - Y_{M4})^2} - d_1$$

Then replace  $d_1$  in  $m_2$ ,  $m_3$ ,  $m_4$  equations above with equivalent  $d_1$  expression from mathematical model to give:

$$m_2 = \sqrt{(X_P - X_{M2})^2 + (Y_P - Y_{M2})^2} - \sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2}$$

$$m_3 = \sqrt{(X_P - X_{M3})^2 + (Y_P - Y_{M3})^2} - \sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2}$$

$$m_4 = \sqrt{(X_P - X_{M4})^2 + (Y_P - Y_{M4})^2} - \sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2}$$

Re-write above three mathematical model equations as *observation equations* by adding a residual  $v_m$  to each measurement:

$$F: m_2 + v_{m2} = \sqrt{(X_P - X_{M2})^2 + (Y_P - Y_{M2})^2} - \sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2}$$

$$G: m_3 + v_{m3} = \sqrt{(X_P - X_{M3})^2 + (Y_P - Y_{M3})^2} - \sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2}$$

$$H: m_4 + v_{m4} = \sqrt{(X_P - X_{M4})^2 + (Y_P - Y_{M4})^2} - \sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2}$$

Because number of measurements ( $m = 3$ ) is greater than number of unknowns ( $n = 2$ ), use Least Squares to determine the MPV of the unknowns  $(X_P, Y_P)$ . Since the observation equations are non-linear in the unknowns  $(X_P, Y_P)$ , a first-order *Taylor Series* is needed to approximate a set of linear observation equations before taking partial derivatives.

Considering function F above (describing ultrasonic relationship between M2 and P):

$$F: m_2 + v_{m2} = \sqrt{(X_P - X_{M2})^2 + (Y_P - Y_{M2})^2} - \sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2}$$

This non-linear function can be written as:

$$F(X_P, Y_P) = m_2 + v_{m2}$$

Where

$$F(X_P, Y_P) = \sqrt{(X_P - X_{M2})^2 + (Y_P - Y_{M2})^2} - \sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2}$$

The above function is *linearized* using a first-order Taylor Series approximation:

$$F(X_p, Y_p) = F(X_{p_o}, Y_{p_o}) + \left( \frac{\partial F}{\partial X_p} \right)_o dX_p + \left( \frac{\partial F}{\partial Y_p} \right)_o dY_p$$

Where

- $X_{p_o}$  and  $Y_{p_o}$  are initial estimates of smartphone position in the room calculated by taking average of all known microphone positions.
- $F(X_{p_o}, Y_{p_o})$  is the non-linear function evaluated with these estimates.
- $dX_p$  and  $dY_p$  are corrections to the initial estimates such that  $X_p = X_{p_o} + dX_p$  and  $Y_p = Y_{p_o} + dY_p$

The partial derivatives  $\left( \frac{\partial F}{\partial X_p} \right)$  and  $\left( \frac{\partial F}{\partial Y_p} \right)$  are found by first re-writing function F:

$$F: \quad F(X_p, Y_p) = \left( (X_p - X_{M2})^2 + (Y_p - Y_{M2})^2 \right)^{\frac{1}{2}} - \left( (X_p - X_{M1})^2 + (Y_p - Y_{M1})^2 \right)^{\frac{1}{2}}$$

and then take partial derivative with respect to  $X_p$ :

$$\begin{aligned} \frac{\partial F}{\partial X_p} &= \frac{1}{2} \left( (X_p - X_{M2})^2 + (Y_p - Y_{M2})^2 \right)^{-\frac{1}{2}} \bullet 2(X_p - X_{M2}) \\ &\quad - \frac{1}{2} \left( (X_p - X_{M1})^2 + (Y_p - Y_{M1})^2 \right)^{-\frac{1}{2}} \bullet 2(X_p - X_{M1}) \\ &= \frac{(X_p - X_{M2})}{\sqrt{(X_p - X_{M2})^2 + (Y_p - Y_{M2})^2}} - \frac{(X_p - X_{M1})}{\sqrt{(X_p - X_{M1})^2 + (Y_p - Y_{M1})^2}} \\ &= \frac{(X_p - X_{M2})}{d_1 + m_2} - \frac{(X_p - X_{M1})}{d_1} \end{aligned}$$

and then with respect to  $Y_P$ :

$$\begin{aligned}\frac{\partial F}{\partial Y_P} &= \frac{1}{2} \left( (X_P - X_{M2})^2 + (Y_P - Y_{M2})^2 \right)^{-\frac{1}{2}} \bullet 2(Y_P - Y_{M2}) \\ &\quad - \frac{1}{2} \left( (X_P - X_{M1})^2 + (Y_P - Y_{M1})^2 \right)^{-\frac{1}{2}} \bullet 2(Y_P - Y_{M1}) \\ &= \frac{(Y_P - Y_{M2})}{\sqrt{(X_P - X_{M2})^2 + (Y_P - Y_{M2})^2}} - \frac{(Y_P - Y_{M1})}{\sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2}} \\ &= \frac{(Y_P - Y_{M2})}{d_1 + m_2} - \frac{(Y_P - Y_{M1})}{d_1}\end{aligned}$$

Where  $d_1$  is always (re)evaluated using Pythagoras at current estimates for  $(X_P, Y_P)$ .

Therefore:

$$\begin{aligned}F(X_P, Y_P) &= F(X_{P_o}, Y_{P_o}) + \left( \frac{(X_P - X_{M2})}{d_1 + m_2} - \frac{(X_P - X_{M1})}{d_1} \right)_o dX_P \\ &\quad + \left( \frac{(Y_P - Y_{M2})}{d_1 + m_2} - \frac{(Y_P - Y_{M1})}{d_1} \right)_o dY_P\end{aligned}$$

So the linearized observation equation for  $m_2$ , describing the ultrasonic relationship between microphone M2 and phone P becomes:

$$\begin{aligned}& \left( \frac{(X_P - X_{M2})}{d_1 + m_2} - \frac{(X_P - X_{M1})}{d_1} \right)_o dX_P + \left( \frac{(Y_P - Y_{M2})}{d_1 + m_2} - \frac{(Y_P - Y_{M1})}{d_1} \right)_o dY_P \\ &= (m_2 - m_{2_o}) + v_{m2}\end{aligned}$$

Likewise for function G (between M3 and P):

$$\left( \frac{(X_P - X_{M3})}{d_1 + m_3} - \frac{(X_P - X_{M1})}{d_1} \right)_o dX_P + \left( \frac{(Y_P - Y_{M3})}{d_1 + m_3} - \frac{(Y_P - Y_{M1})}{d_1} \right)_o dY_P \\ = (m_3 - m_{3o}) + v_{m3}$$

and function H (between M4 and P):

$$\left( \frac{(X_P - X_{M4})}{d_1 + m_4} - \frac{(X_P - X_{M1})}{d_1} \right)_o dX_P + \left( \frac{(Y_P - Y_{M4})}{d_1 + m_4} - \frac{(Y_P - Y_{M1})}{d_1} \right)_o dY_P \\ = (m_4 - m_{4o}) + v_{m4}$$

When using Matrix Methods for Least Squares, the observation equations are represented in matrix form as:

$${}_m A_{n\ n} X_1 = {}_m L_1 + {}_m V_1$$

Where in our case:

- $m = 3, n = 2$
- ${}_m A_n$  contains the coefficients of the unknowns  $(X_P, Y_P)$
- ${}_n X_1$  contains the corrections to be applied to the initial estimates for the unknowns  $(dX_P, dY_P)$
- ${}_m L_1$  contains the measurements  $(m_2, m_3, m_4)$
- ${}_m V_1$  contains the residuals (one for each measurement)

Solving for  $X$  gives the solution:

$$X = (A^T A)^{-1} A^T L \text{ where:} \quad (4-5)$$

$$A = \begin{bmatrix} \frac{(X_P - X_{M2})}{d_1 + m_2} - \frac{(X_P - X_{M1})}{d_1} & \frac{(Y_P - Y_{M2})}{d_1 + m_2} - \frac{(Y_P - Y_{M1})}{d_1} \\ \frac{(X_P - X_{M3})}{d_1 + m_3} - \frac{(X_P - X_{M1})}{d_1} & \frac{(Y_P - Y_{M3})}{d_1 + m_3} - \frac{(Y_P - Y_{M1})}{d_1} \\ \frac{(X_P - X_{M4})}{d_1 + m_4} - \frac{(X_P - X_{M1})}{d_1} & \frac{(Y_P - Y_{M4})}{d_1 + m_4} - \frac{(Y_P - Y_{M1})}{d_1} \end{bmatrix} \quad (4-6)$$

$$X = \begin{bmatrix} dX_P \\ dY_P \end{bmatrix} \quad (4-7)$$

$$L = \begin{bmatrix} m_2 - m_{2_0} \\ m_3 - m_{3_0} \\ m_4 - m_{4_0} \end{bmatrix} \quad (4-8)$$

$$V = \begin{bmatrix} v_{m2} \\ v_{m3} \\ v_{m4} \end{bmatrix} \quad (4-9)$$

Matrix  $X$  contains the corrections to be applied to the original estimates for  $(X_P, Y_P)$ .

These new  $(X_P, Y_P)$  coordinates are then used to recalculate updated distances for  $(d_1, m_{2_0}, m_{3_0}, m_{4_0})$ . The process is repeated until coordinates of  $(X_P, Y_P)$  don't change significantly (e.g. in the 3<sup>rd</sup> decimal place for mm precision). For the initial estimation  $X_p$  and  $Y_p$  can be set to the average of x and y coordinates of the four microphones:

$$X_p = \frac{X_{M1} + X_{M2} + X_{M3} + X_{M4}}{4} \quad (4-10)$$

and

$$Y_p = \frac{Y_{M1} + Y_{M2} + Y_{M3} + Y_{M4}}{4} \quad (4-11)$$

$d_1$  can be estimated from the  $X_p$  and  $Y_p$  from Pythagoras:

$$d_1 = \sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2} \quad (4-12)$$

similarly  $(m_{2_0}, m_{3_0}, m_{4_0})$  can be calculated:

$$m_{2_0} = \sqrt{(X_P - X_{M2})^2 + (Y_P - Y_{M2})^2} - d_1 \quad (4-13)$$

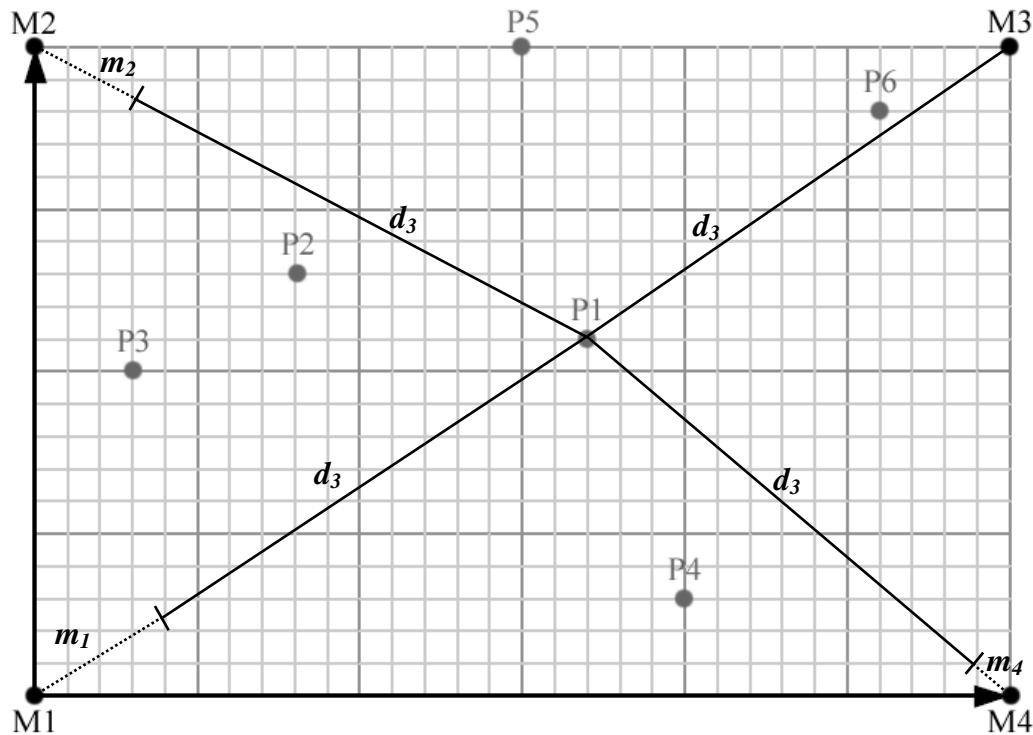
$$m_{3_0} = \sqrt{(X_P - X_{M3})^2 + (Y_P - Y_{M3})^2} - d_1 \quad (4-14)$$

$$m_{4_0} = \sqrt{(X_P - X_{M4})^2 + (Y_P - Y_{M4})^2} - d_1 \quad (4-15)$$

The difference between matrices in **TOA** and **TDOA** solutions can be described and remembered as follows. In **TDOA** the first row dedicated to point 1 is missing from matrices L (4-8) and V (4-9), unlike in **TOA** (4-3),(4-4). In order to transform **TOA** A matrix (4-1) to **TDOA** A matrix (4-6), variable  $d_1$  is added to the divisor in each element of A matrix, with the exception of the first row where  $d_1$  replaces the divisor entirely. After that the first element of the first column is removed and subtracted from each of the elements in the first column. Similarly the first element of the second column is removed and subtracted from each element of the second column. This completes the transition from **TOA** A matrix (4-1) to **TDOA** A matrix (4-6). Matrix X is the same in **TOA** (4-2) and **TDOA** (4-7) solutions. This matrix manipulation is what allows for extending standard Least Squares trilateration to the case where distances between unknown target position and known control points are unknown a-priori.

### 4.3. Working Example of TDOA Trilateration

To test the accuracy of our TDOA Trilateration method, we used it to calculate the position of several random smartphone locations and compare the results to their actual positions in Figure 30. We used four control points (microphones) arranged in the corners of a rectangular room to locate the phone's position at 6 different locations within the room.



**Figure 30: TDOA Trilateration experiment with four microphones and six different smartphone positions.** Control points  $M_1$ ,  $M_2$ ,  $M_3$  and  $M_4$  are microphones. Points  $P_1$ ,  $P_2$ ,  $P_3$ ,  $P_4$ ,  $P_5$  and  $P_6$  are actual smartphone locations. Each square of the grid represents 1 unit in length.

Regarding input data for testing the Lok8 TDOA trilateration algorithm, the locations of  $M_1(0,0)$ ,  $M_2(0,20)$ ,  $M_3(30,20)$ ,  $M_4(30,0)$  were used and the initial distances between the mics and the various phone positions were measured manually. Although we could have used Pythagoras in Figure 30 to calculate exactly the measurements representing the ultrasonic distances between the microphones and various phone positions on the

grid, we wanted to introduce some error in the measurements so chose instead to simply use a ruler to measure these distances to one decimal point precision. After that we subtracted the shortest measured distance for any given phone position from each of the remaining three mic distances. The resulting 3 distance differences were then used as “ultrasonic” input in addition to the known microphone locations.

For example, for phone position P1 the measured distance to M1 is 20.2, to M2 19.2, M3 15.8, and M4 17.0. The shortest distance is to M3, therefore it is subtracted from the other 3 distances to leave;  $m_1 = 4.4$ ,  $m_2 = 3.4$ ,  $m_4 = 1.2$ . These values simulate time measurements translated to distance for the ultrasonic signal to reach these 3 mics after first triggering the server clock at M3. Since in the solution M1 is supposed to be the closest microphone, the microphone names are switched around to account for that. M3 becomes M1 and M1 becomes M3. The exact order in which microphones are numbered is not important as long as M1 is the closest. The input data is summarised in Table 3. For the first iteration  $X_p$  and  $Y_p$  are calculated using Equation (4-10) and (4-11).

$$X_p = \frac{X_{M1} + X_{M2} + X_{M3} + X_{M4}}{4} = \frac{30 + 0 + 0 + 30}{4} = 15$$

$$Y_p = \frac{Y_{M1} + Y_{M2} + Y_{M3} + Y_{M4}}{4} = \frac{20 + 20 + 0 + 0}{4} = 10$$

After this  $d_1$  is calculated using Equation (4-12):

$$d_1 = \sqrt{(X_p - X_{M1})^2 + (Y_p - Y_{M1})^2} = \sqrt{(15 - 30)^2 + (10 - 20)^2} = 18.0277563773$$

$(m_{2_0}, m_{3_0}, m_{4_0})$  are also estimated using Equation (4-13), (4-14), (4-15):

$$m_{2_0} = \sqrt{(X_P - X_{M2})^2 + (Y_P - Y_{M2})^2} = \sqrt{(15 - 0)^2 + (10 - 20)^2} - 18.0277563773 = 0$$

$$m_{3_0} = \sqrt{(X_P - X_{M3})^2 + (Y_P - Y_{M3})^2} = \sqrt{(15 - 0)^2 + (10 - 20)^2} - 18.0277563773 = 0$$

$$m_{4_0} = \sqrt{(X_P - X_{M4})^2 + (Y_P - Y_{M4})^2} = \sqrt{(15 - 30)^2 + (10 - 20)^2} - 18.0277563773 = 0$$

For the first iteration  $(m_{2_0}, m_{3_0}, m_{4_0})$  will have the value zero because initial values for  $X_P$  and  $Y_P$  are at equal distance from all four microphones. After this A and L matrices are populated as in Equations (4-6) and (4-8).

$$\begin{aligned} A &= \begin{bmatrix} \frac{(X_P - X_{M2})}{d_1 + m_2} - \frac{(X_P - X_{M1})}{d_1} & \frac{(Y_P - Y_{M2})}{d_1 + m_2} - \frac{(Y_P - Y_{M1})}{d_1} \\ \frac{(X_P - X_{M3})}{d_1 + m_3} - \frac{(X_P - X_{M1})}{d_1} & \frac{(Y_P - Y_{M3})}{d_1 + m_3} - \frac{(Y_P - Y_{M1})}{d_1} \\ \frac{(X_P - X_{M4})}{d_1 + m_4} - \frac{(X_P - X_{M1})}{d_1} & \frac{(Y_P - Y_{M4})}{d_1 + m_4} - \frac{(Y_P - Y_{M1})}{d_1} \end{bmatrix} \\ &= \begin{bmatrix} \frac{(15 - 0)}{18.027756 + 3.4} - \frac{(15 - 30)}{18.027756} & \frac{(10 - 20)}{18.027756 + 3.4} - \frac{(10 - 20)}{18.027756} \\ \frac{(15 - 0)}{18.027756 + 4.4} - \frac{(15 - 30)}{18.027756} & \frac{(10 - 0)}{18.027756 + 4.4} - \frac{(10 - 20)}{18.027756} \\ \frac{(15 - 30)}{18.027756 + 1.2} - \frac{(15 - 30)}{18.027756} & \frac{(10 - 0)}{18.027756 + 1.2} - \frac{(10 - 20)}{18.027756} \end{bmatrix} \\ &= \begin{bmatrix} 1.5320769 & 0.08801578 \\ 1.5008644 & 1.00057627 \\ 0.051928 & 1.07478168 \end{bmatrix} \end{aligned}$$

$$L = \begin{bmatrix} m_2 - m_{2_0} \\ m_3 - m_{3_0} \\ m_4 - m_{4_0} \end{bmatrix} = \begin{bmatrix} 3.4 - 0 \\ 4.4 - 0 \\ 1.2 - 0 \end{bmatrix} = \begin{bmatrix} 3.4 \\ 4.4 \\ 1.2 \end{bmatrix}$$

$A^T$  matrix is calculated by transposing A matrix:

$$A^T = \begin{bmatrix} 1.5320769 & 0.08801578 \\ 1.5008644 & 1.00057627 \\ 0.051928 & 1.07478168 \end{bmatrix}^T = \begin{bmatrix} 1.5320769 & 1.5008644 & 0.051928 \\ 0.08801578 & 1.00057627 & 1.07478168 \end{bmatrix}$$

$X$  matrix is calculated using Equation (4-5):

$$\begin{aligned} X = (A^T A)^{-1} A^T L &= \left( \begin{bmatrix} 1.5320769 & 1.5008644 & 0.051928 \\ 0.08801578 & 1.00057627 & 1.07478168 \end{bmatrix} \bullet \begin{bmatrix} 1.5320769 & 0.08801578 \\ 1.5008644 & 1.00057627 \\ 0.051928 & 1.07478168 \end{bmatrix} \right)^{-1} \\ &\bullet \begin{bmatrix} 1.5320769 & 1.5008644 & 0.051928 \\ 0.08801578 & 1.00057627 & 1.07478168 \end{bmatrix} \bullet \begin{bmatrix} 3.4 \\ 4.4 \\ 1.2 \end{bmatrix} \\ &= \begin{bmatrix} 4.6025501855575355 & 1.6923875995963158 \\ 1.6923875995963158 & 2.164055305489902 \end{bmatrix}^{-1} \bullet \begin{bmatrix} 11.875178612467845 \\ 5.991527256460969 \end{bmatrix} \\ &= \begin{bmatrix} 2.192582107615011 \\ 1.0539603496502288 \end{bmatrix} \end{aligned}$$

$X$  matrix contains corrections to be applied to  $X_p$  and  $Y_p$ . This process is repeated until the corrections become sufficiently low. At that point  $X_p$  and  $Y_p$  should contain the correct position of the phone.

**Table 3: Sample TDOA Trilateration input.** Second and third columns contain coordinates of a microphone and fourth column contains differences between distance to mic and closest mic. In this example microphone M3 is closest to phone position P1 so it has been switched around with M1.

Mic	X	Y	Distance Difference ( $m_i$ )
M1	30	20	0
M2	0	20	3.4
M3	0	0	4.4
M4	30	0	1.2

Trilateration results for the phone's position at P1-P6 are compiled in Table 4. Notice that if we assumed metres for units in this example, the standard deviations for the phone positions are of sub-metre accuracy after only a few iterations.

**Table 4: Comparison of TDOA output and expected results.** Second column contains actual X and Y coordinates of a given phone position, third column contains coordinates of the phone as calculated by our TDOA trilateration procedure. Fourth and fifth columns contain the Standard Deviations ( $\sigma_x, \sigma_y$ ) for each trilaterated phone positon and number of iterations to get there.

Phone Point	Actual Location	TDOA Trilateration	Standard Deviation	Number of Iterations
P1	17 , 11	16.987, 10.986	0.0002, 0.0003	3
P2	8 , 13	7.978, 12.966	0.0158, 0.019	3
P3	3 , 10	2.96, 10.0	0 , 0	4
P4	20 , 3	20.002, 2.996	0.011, 0.0195	3
P5	15 , 20	15.0, 20.0	0 , 0	4
P6	26 , 18	25.999, 18.031	0.0144, 0.0214	4

## 5. EVALUATION

This chapter is dedicated to evaluating our indoor positioning approach in a real-world environment, finding its strengths and weaknesses and determining its accuracy. This is done through a range of experiments that involve variables such as position of the microphones, position of the phone in relation to the centre of the room, direction the speaker is facing, presence of the user in the path of the signal and background noise.

### 5.1. Positioning Accuracy

This section addresses the fifth research question (RQ5): What accuracy can mobile asynchronous ultrasound trilateration offer? It also seeks to combine and utilise findings from previous experiments.

The aim of the experiments described in this section is to find out with what accuracy a mobile phone emitting an ultrasound signal can be positioned using four microphones placed in every corner of a 7 by 7 metre room. The signal is the same as depicted in Figure 24, Section 3.2. The positioning approach described in the previous chapter assumes that the phone and the microphones are located roughly on the same plane. In most scenarios this is impractical. A more suitable placement for microphones would be at ceiling level. There are a number of advantages.

- Office furniture and other obstacles are less likely to block the direct line-of-sight.
- Microphones will occupy space that is otherwise not used.
- Microphones are less likely to be interfered with.
- It is easier to wire microphones provided most offices have a dropped ceiling.

Therefore three experiments were carried out. During Experiment 1 the microphones and the phone were positioned at the same height 0.12 m below the ceiling. This experiment closely follows the scenario given in Section 3.4 where the phone and microphones are all in the same plane. During Experiment 2 the phone was lowered to chest height putting it roughly 1.6 m below microphone level. No adjustments were made to the trilateration algorithms compared to Experiment 1. During Experiment 3 the phone remained at the same height as in Experiment 2, but a “room calibration factor” was introduced into the trilateration algorithms in order to decrease any impact of height difference on accuracy.

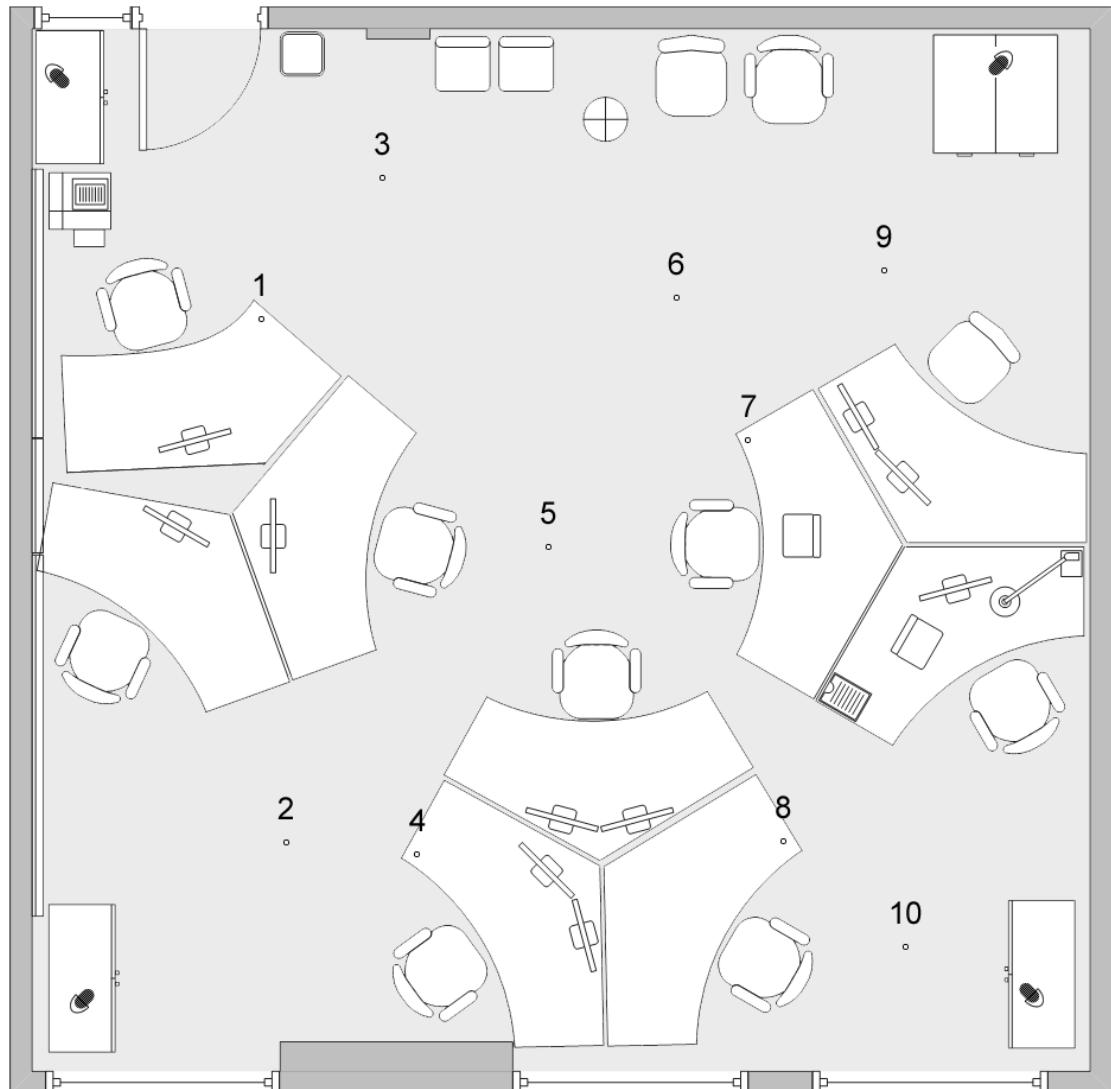
**Experiment Method.** Positioning was done in an office room 7.1 m long, 7 m wide and 2.83 m high. Layout of the room can be seen on Figure 31. Four DPA microphones were placed in every corner of the room 10-20 cm below the ceiling, facing the lower opposite corner of the room. All microphones were connected to Avid Mbox Pro audiocard, which supports four-channel synchronised input. The audiocard was in turn connected to a PC running LOK8 asynchronous trilateration software.

The software uses RtAudio API for real-time audio input/output. It is accessed via a Java wrapper JRtAudio, which allowed quick GUI prototyping in Java. On PC RtAudio API supports either DirectSound or ASIO drivers. ASIO support proved essential to us, because streaming more than two channels with DirectSound turned out to be very challenging. Audio streams from the audiocard are written into a buffer and get processed through a bandpass filter. We found the Bessel bandpass filter (Paarmann 2001) with filter order 4, corner frequency one = 21400 Hz and corner frequency two = 21600 Hz to have the best response to the phone’s ultrasonic signal. This filters out all the other frequencies and leaves almost only the frequency of the signal. Sound is

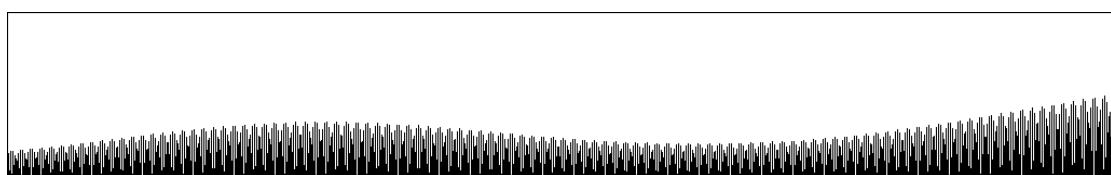
streamed at 96000 bit/s sampling rate. Assuming speed of sound is 346 m/s, each sample should be equivalent to 3.6 mm travelled. Unfortunately not every sample of the bandpass filter is a valid representation of the signal's intensity. Figure 32 shows a sample filter output. It can be seen that the graph entirely consists of groups of 10 samples. These groups appear regularly and carry no useful information other than the height of the loudest sample in the middle. The other 9 samples can be discarded, being a by-product of using a low-order filter. Therefore the real resolution of the system is 3.6 multiplied by 10 or **36 mm**. This parameter can be improved by using a high-order filter, which would require more powerful hardware to run in real-time or even the use of a Digital Signal Processor (DSP).

The same ten random test locations or check points in the room were chosen for all three experiments. Precise coordinates of these locations were measured with a laser measurement device. These test locations can be seen on Figure 33. The phone was placed at each of these locations with its main speaker pointing directly upward. During Experiment 1 the phone was 0.16 m below the ceiling and during Experiment 2 & 3 it was 1.7 m below the ceiling to mimic normal carrying height. For Experiment 3 the positioning software was altered to accommodate multiplication by a room calibration factor of 1.1 for each difference in delay. This value appears to compensate well for the height difference between microphones and the phone and was found by trial and error. During all three experiments the phone was made to produce an ultrasonic pulse 100 times at each of the 10 locations with an interval of one second between pulses to provide enough measurement data for statistical analysis. These signals were captured by the four microphones and processed by LOK8 positioning software that independently estimated coordinates of the signal source for each detected signal. These coordinates as well as an estimated standard deviation for the trilateration procedure

were recorded into text files. This resulted in 3000 readings; 300 per each unique location, 1000 per experiment.



**Figure 31: Layout of the room used for trilateration accuracy experiments.** Dimensions of the room are 7.1 by 7 meters. The position of 10 test locations or check points is shown as well as the position of the four microphones placed at near ceiling height in each corner.



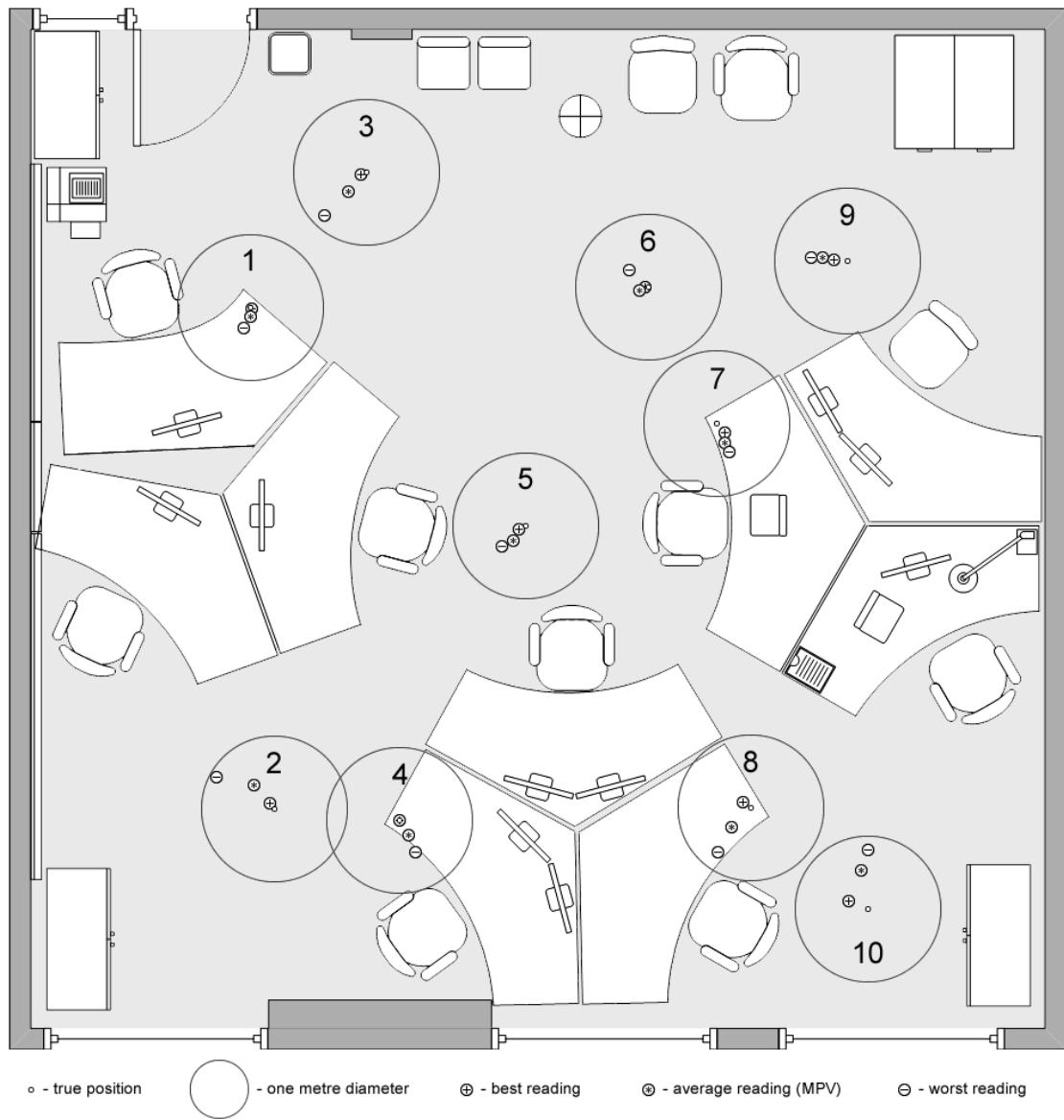
**Figure 32: Raw filter output.** The height of each column represents intensity. It was observed that every 1 out of 10 samples can be used to reliably estimate intensity of a given frequency.

**Table 5: Difference between MPV and true position  $\pm$  standard deviation.**

*First column contains the check point number, which are used to identify each of the ten known test positions in the rest of the chapter. Columns 2,3 and 4 contain error and standard deviation for each test in each of the three experiments.*

Check point number	Experiment 1 error $\pm$ standard deviation (m)	Experiment 2 error $\pm$ standard deviation (m)	Experiment 3 error $\pm$ standard deviation (m)
1	$0.072 \pm 0.051$	$0.172 \pm 0.047$	$0.070 \pm 0.043$
2	$0.219 \pm 0.097$	$0.236 \pm 0.063$	$0.081 \pm 0.035$
3	$0.184 \pm 0.076$	$0.257 \pm 0.048$	$0.083 \pm 0.046$
4	$0.121 \pm 0.053$	$0.122 \pm 0.049$	$0.086 \pm 0.039$
5	$0.140 \pm 0.057$	$0.144 \pm 0.057$	$0.133 \pm 0.028$
6	$0.064 \pm 0.056$	$0.166 \pm 0.048$	$0.080 \pm 0.030$
7	$0.151 \pm 0.042$	$0.089 \pm 0.044$	$0.097 \pm 0.033$
8	$0.187 \pm 0.068$	$0.288 \pm 0.058$	$0.061 \pm 0.039$
9	$0.175 \pm 0.048$	$0.182 \pm 0.078$	$0.033 \pm 0.099$
10	$0.277 \pm 0.082$	$0.153 \pm 0.142$	$0.099 \pm 0.041$
<b>Average</b>	<b><math>0.159 \pm 0.063</math></b>	<b><math>0.181 \pm 0.063</math></b>	<b><math>0.082 \pm 0.043</math></b>

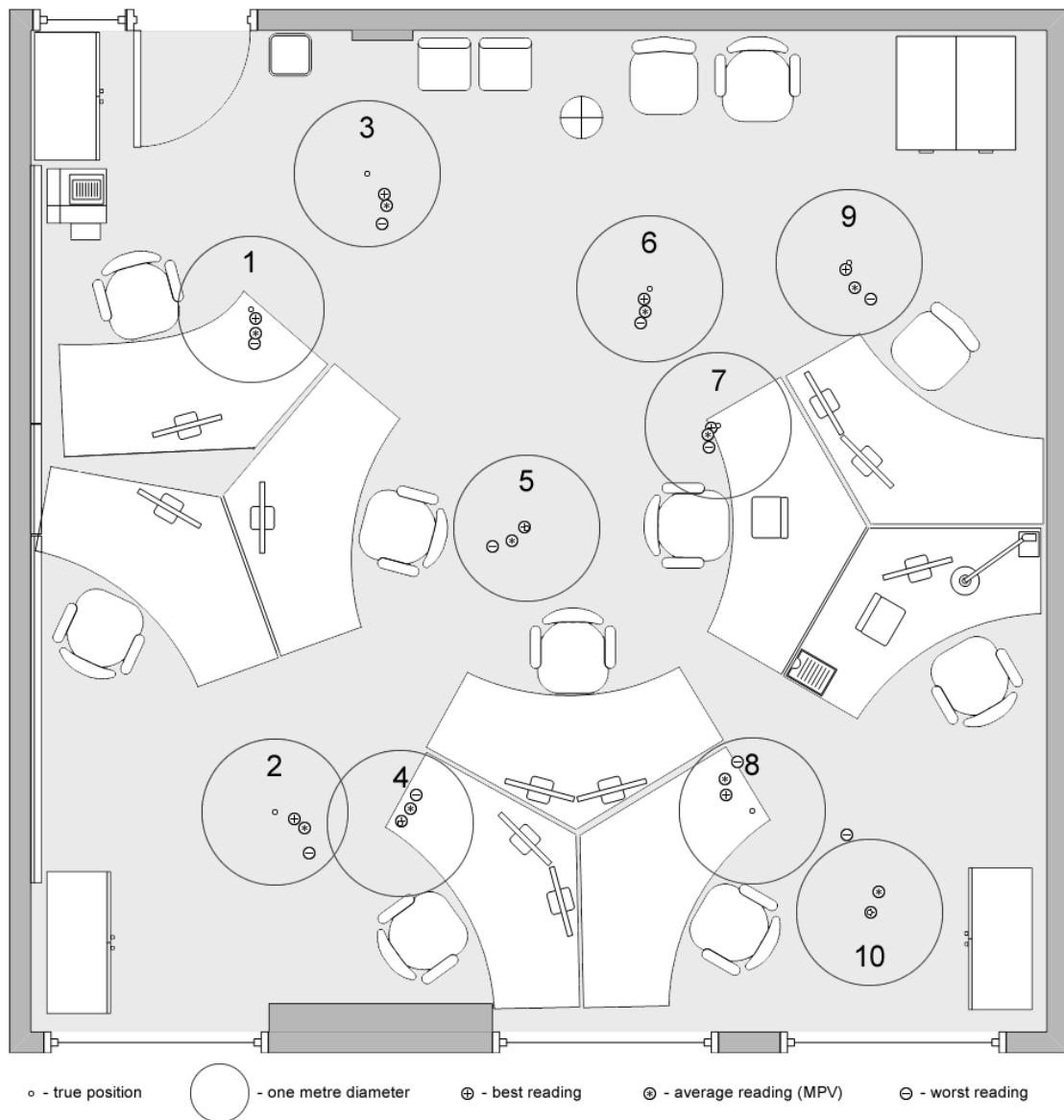
**Discussion.** All one hundred readings collected at each unique test location during each of the three experiments were treated as one sample. For each sample the Most Probable Value (MPV) was calculated. The difference between MPV and true position of the check point as well as its standard deviation can be found in Table 5. Also for each sample, average, best and worst results were marked on the floor plan. See Figures 33-35. Coordinates of average, best and worst results as well as the true position for all check points can be found in Appendix 4.



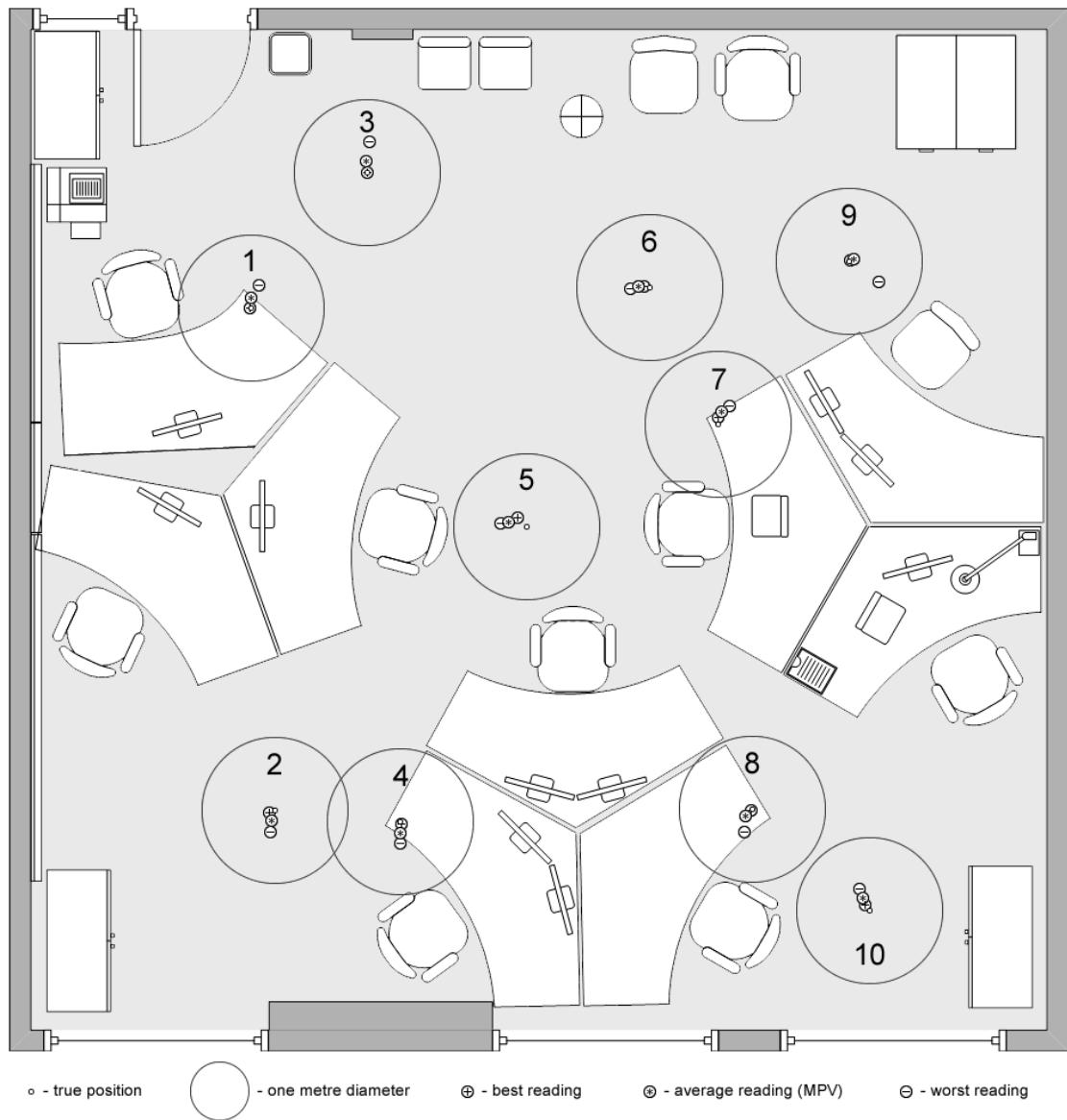
**Figure 33: Average, best and worst results for Experiment 1 (2D Trilateration).**

On average, Experiment 3 (3D trilateration with room calibration factor) showed the best accuracy and experiment 2 (3D trilateration without calibration) the worst. Average difference in accuracy between Experiments 1 (2D trilateration) & 2 is only around 14%, whereas 3 has twice better accuracy than 1 and more than twice better accuracy than 2. Also, Experiment 3 on average has a smaller standard deviation. Interestingly, test point 5, located in the middle of the room, gave approximately the same error in all three experiments. It was the worst result among Experiment 3 and among the best results in Experiment 1 & 2. In general, all three experiments showed stable sub-metre

accuracy. Only in Experiment 2, test point 10, the worst reading was more than half a metre away from the true position, although when averaged with the other readings at this point produced an error of only 15.3 cm. Experiment 3 on average had sub-decimetre accuracy.



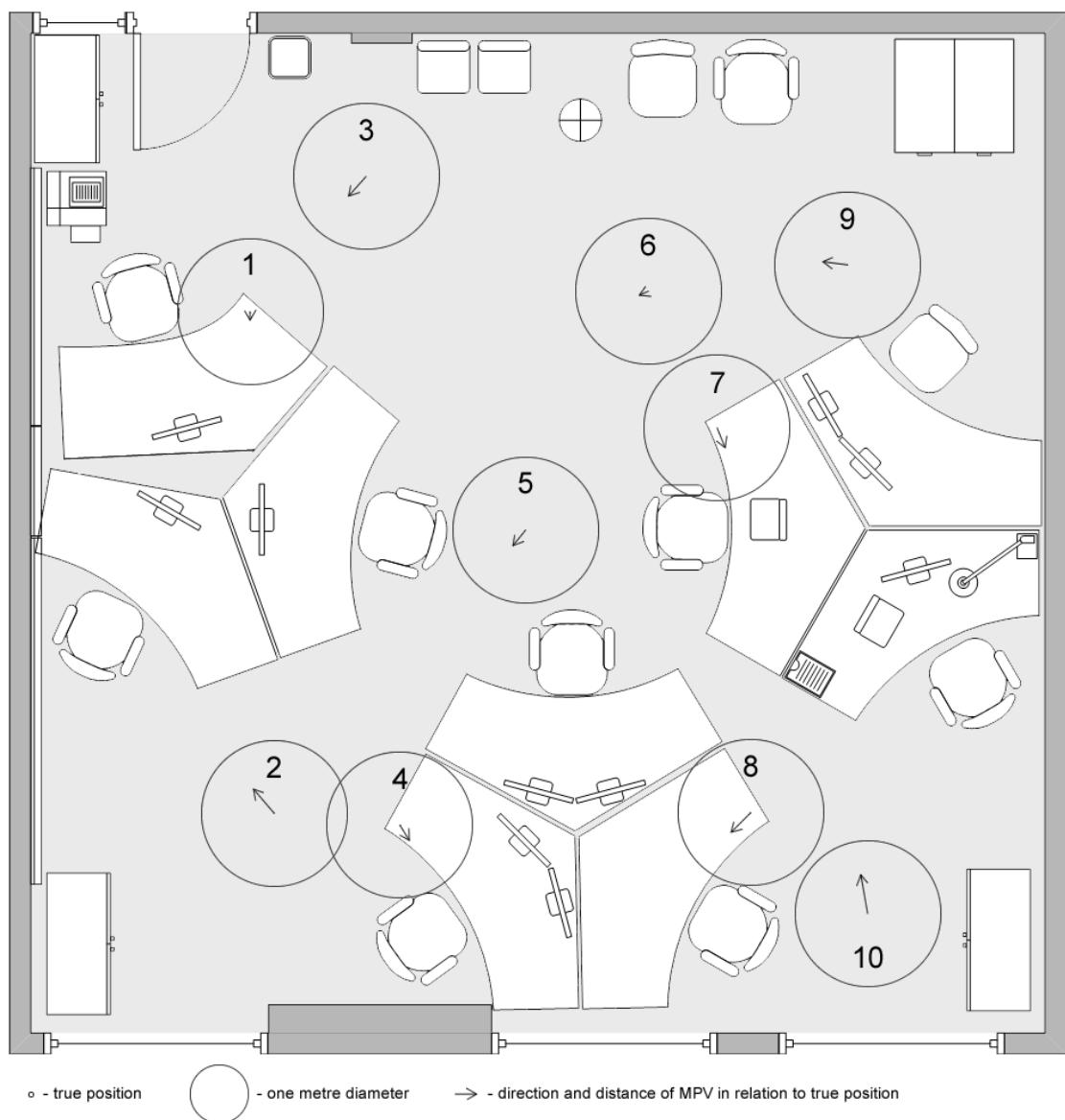
**Figure 34: Average, best and worst results for Experiment 2 (3D Trilateration).**



**Figure 35: Average, best and worst results for Experiment 3 (3D Trilateration with room calibration factor of 1.1).**

In order to analyze whether there is a particular direction/pattern in which trilateration calculations (on average) were displaced in relation to true position, arrows indicating distance and direction of error displacement were placed on the floor plan for each of the three experiments (See Figures 36-38). In the first experiment there is no clear trend in which the error displacement occurs, so it is unlikely to have been caused by systematic error created by incorrect calibration or measurements. When compared to Experiment 3, the error is relatively large. Most likely this is the direct result of the

mobile phone and microphones being positioned on the same plane. The speaker on the phone was pointing directly upward, so being on the same plane meant the source was rotated roughly 90 degrees in relation to each microphone. Given the highly directional nature of ultrasound this resulted in a weaker, less stable signal detection even though the average distance to each microphone was shorter. In the other two experiments, the phone was roughly 1.6 m lower than the microphones, resulting in the source being positioned at an angle substantially less than 90 degrees to the microphones.

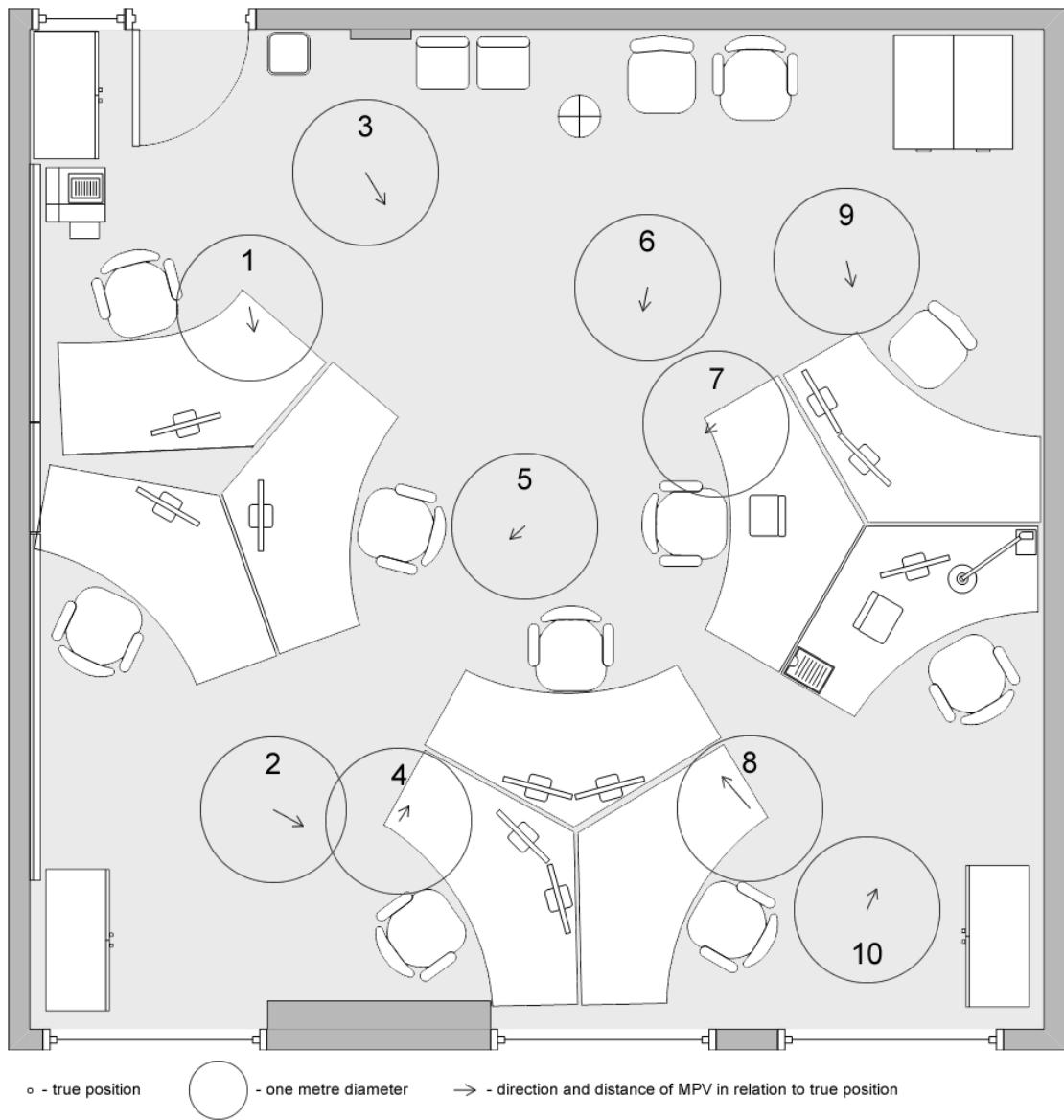


**Figure 36: Direction and distance from true position to MPV in Experiment 1 (2D Trilateration).**

For Experiment 2, it can be observed that arrows mostly point inward, towards the centre of the room. This is a predictable outcome. In Experiment 2 & 3, delay is calculated based on signals that arrived via direct line-of-sight, whereas the 2D position of the phone is calculated based on the projection of these direct lines-of-sight to a 2D plane. Bigger height difference between microphones and the phone generally means reduced difference in the time of arrival, causing readings to appear closer to the centre of the room. Unfortunately, the relationship between distance to microphones and delay reduction is not linear, making it necessary to know a phone's position before the required compensation can be calculated. Since in our case the phone's position is unknown before trilateration is done, it becomes necessary to take calculations required to make a projection of direct lines of sight onto a plane and incorporate them into the asynchronous trilateration procedure. This is a complex task and is left for future work, although overall positioning accuracy is still sub-metre without introduction of this correction. In Experiment 3 a simpler constant multiplier (room calibration factor) was used instead.

In Experiment 3 each delay was multiplied by a constant value of 1.1 (room calibration factor), which produced significant improvements in accuracy over Experiment 2 from 18 cm to 8 cm. Although the use of a constant multiplier may in some cases result in a decrease in accuracy, it appears to affect accuracy relatively less when compared to other factors. The biggest factor appears to be poor signal reception when the phone is in the corners or along the walls. Usually this happens if the microphone is located far away and the direction it is pointing at is substantially different from the direction in which the phone is located. One possible solution to this problem could be the use of microphones that are less directional. Reception in these areas was also influenced by yaw-orientation of the phone. This was an unexpected factor. Theoretically, since the speaker points directly upwards, yaw-orientation should have no influence on the

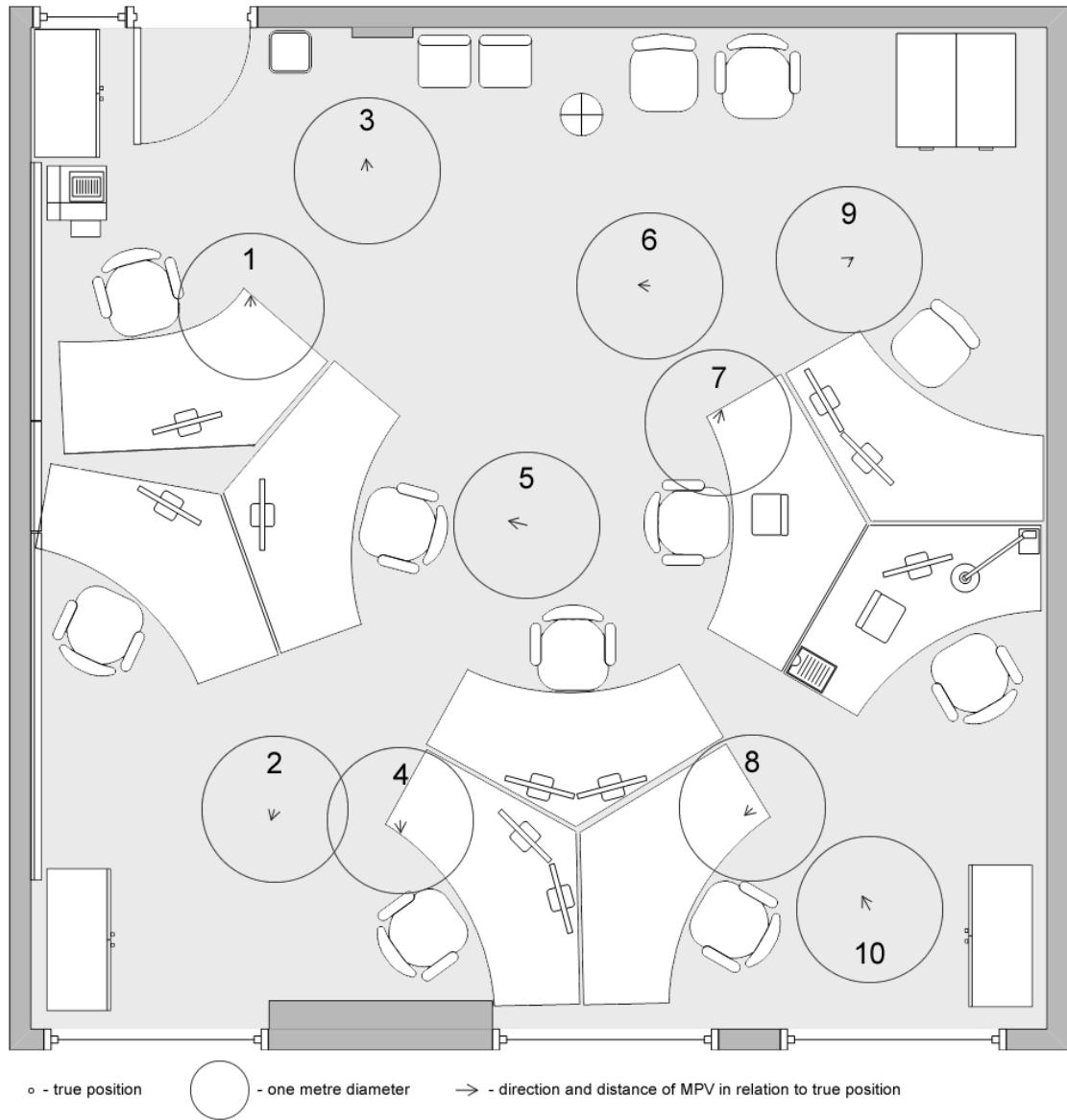
strength of the signal upon arrival, unless the phone's speaker is also somewhat directional in its manufacture.



**Figure 37: Direction and distance from true position to MPV in Experiment 2 (3D Trilateration).**

There appears to be no correlation between the standard deviation, which was calculated during trilateration for each individual reading, and its proximity either to the true position (absolute value of the correlation coefficient  $<0.04$ ) or MPV (absolute value of the correlation coefficient  $<0.06$ ). This is unfortunate, because we wanted to see whether it is possible to predict how accurate an individual reading is based on the

program's output. It is, however possible to use large standard deviation values ( $>250$  mm) to filter out false readings caused by something other than the signal.



**Figure 38: Direction and distance from true position to MPV in Experiment 3 (3D Trilateration with room calibration factor of 1.1).**

## 5.2. Angle Variation

This section addresses the sixth research question (RQ6): What impact can orientation of the speaker and the way user stands have on accuracy and reliability? In the previous

section all experiments were done with the phone's speaker facing upwards. This orientation should be optimal for maximising signal reception if microphones were placed just below the ceiling in each corner of the room. Because ultrasound is very directional, it is important to minimise the angle between the direction the speaker is facing and the line-of-sight between the speaker and all four microphones. With the speaker pointing upwards this angle will never exceed 90 degrees for all four microphones, which is important for this implementation, where reception at four microphones is the minimum required. Unfortunately the speaker on the front panel of the phone used as an earpiece is not powerful enough. The multimedia speaker used for ringtones and other sounds that are supposed to be heard at distance is normally placed on the back of the phone. However, while the user is interacting with the touchscreen, the speaker will normally be facing downwards. Therefore the aim of this experiment is to test how well the positioning system will work with the phone speaker inclined at various angles other than straight up.

Unless the speaker is pointing directly upwards or downwards, it becomes important what yaw orientation the mobile phone has in addition to its incline (pitch). Another factor is the introduction of the user holding the phone. He may or may not be blocking the direct line-of-sight to one of the four microphones. Because this test setup introduces so many new variables, some of them had to be taken out of the equation. Yaw orientation for each location was chosen in such a way that the speaker is facing the farthest microphone. Also the user tries not to block direct line-of-sight to each of the microphones if possible. The experiment is designed to measure what effects incline (pitch) variation alone can have on positioning accuracy and failure rate. The variables omitted in this experiment are addressed in the next section.

**Experiment Method.** Positioning was done in the same lab room as in the previous section, 7.1m long, 7m wide and 2.83m high. From this point onwards only 3D trilateration with room calibration was used for positioning. The same 10 check points were used as in the previous section (See Figure 31). At each check point the user stood in such a way as not to block direct line-of-sight to any of the microphones. He held the phone in 5 different orientations: speaker pointing downwards (0 degrees), halfway between horizontal and downwards (45 degrees), horizontal (90 degrees), halfway between horizontal and upwards (135 degrees) and upwards (180 degrees). For each of the five orientations 10 readings were made. For each check point, yaw angle was chosen in such a way that the speaker is oriented towards the farthest microphone in order to improve overall signal detection.

**Discussion.** All 10 readings collected at 5 different orientations at each unique check point were treated as one sample. For each sample the Most Probable Value (MPV) was calculated. The difference between MPV and true position of the check point as well as its standard deviation can be found in Table 6. From the average of all 10 check points it appears that accuracy steadily drops as the speaker is rotated from upward to downward orientation. However this trend is not true for every single point individually. See Figure 39.

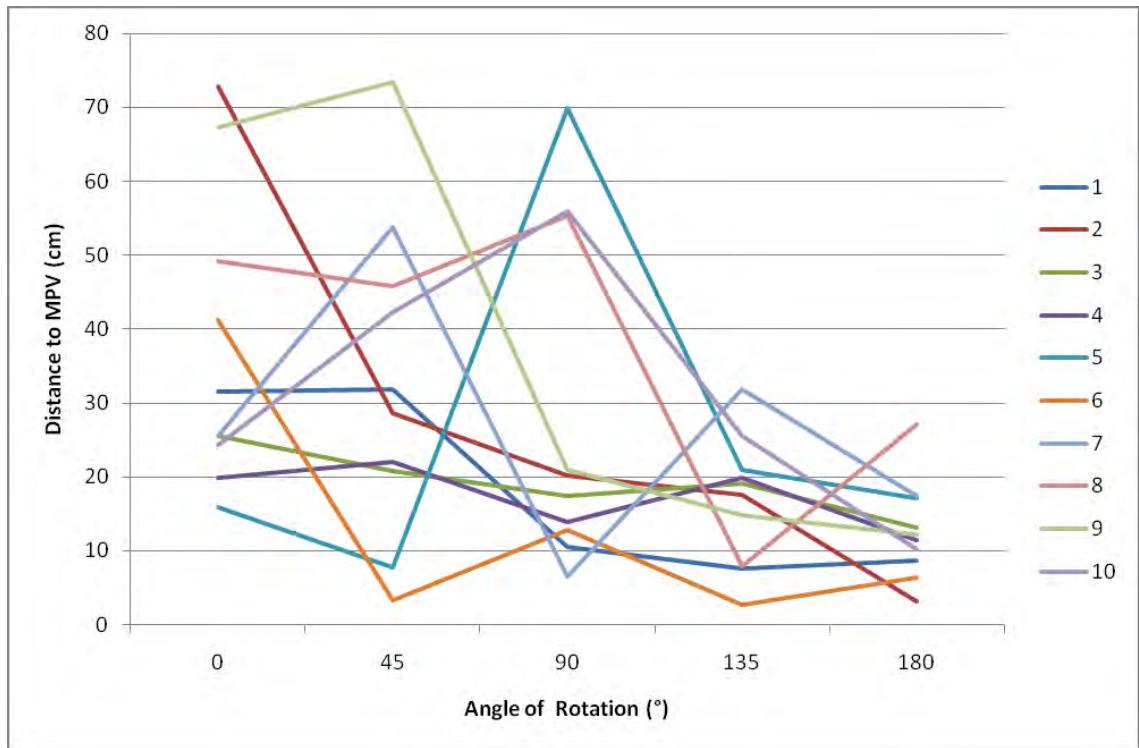
The percentage of detected signals can be found in Table 7. Signal detection was considered to have failed either if the system failed to produce an output or the calculated position was outside the dimensions of the room.

**Table 6: Difference between MPV and true position  $\pm$  standard deviation for different pitch angles.** First column contains the check point number, which are used to identify each of the ten known test positions in the rest of the chapter. Columns 2, 3, 4, 5 and 6 contain error and standard deviation for each angle at which the phone was held. The bottom row is average of the 10 check points.

	0 (downwards)	45	90	135	180 (upwards)
1	$0.315 \pm 0.106$	$0.319 \pm 0.116$	$0.105 \pm 0.062$	$0.076 \pm 0.036$	$0.087 \pm 0.016$
2	$0.728 \pm 0.282$	$0.286 \pm 0.045$	$0.203 \pm 0.045$	$0.176 \pm 0.023$	$0.032 \pm 0.032$
3	$0.255 \pm 0.038$	$0.209 \pm 0.042$	$0.174 \pm 0.026$	$0.191 \pm 0.018$	$0.132 \pm 0.022$
4	$0.199 \pm 0.138$	$0.221 \pm 0.049$	$0.140 \pm 0.121$	$0.200 \pm 0.132$	$0.115 \pm 0.028$
5	$0.159 \pm 0.107$	$0.077 \pm 0.061$	$0.698 \pm 0.040$	$0.209 \pm 0.024$	$0.171 \pm 0.013$
6	$0.412 \pm 0.144$	$0.033 \pm 0.094$	$0.127 \pm 0.022$	$0.026 \pm 0.029$	$0.062 \pm 0.013$
7	$0.256 \pm 0.200$	$0.537 \pm 0.054$	$0.066 \pm 0.049$	$0.319 \pm 0.020$	$0.175 \pm 0.021$
8	$0.492 \pm 0.052$	$0.457 \pm 0.210$	$0.553 \pm 0.472$	$0.079 \pm 0.035$	$0.27 \pm 0.025$
9	$0.674 \pm 0.099$	$0.735 \pm 0.026$	$0.210 \pm 0.035$	$0.147 \pm 0.037$	$0.122 \pm 0.013$
10	$0.243 \pm 0.023$	$0.423 \pm 0.105$	$0.559 \pm 0.139$	$0.255 \pm 0.043$	$0.102 \pm 0.031$
avg	<b><math>0.373 \pm 0.119</math></b>	<b><math>0.330 \pm 0.080</math></b>	<b><math>0.283 \pm 0.101</math></b>	<b><math>0.168 \pm 0.040</math></b>	<b><math>0.127 \pm 0.021</math></b>

**Table 7: Percentage of detected signals for different angles.** First column contains the check point number. Columns 2, 3, 4, 5 and 6 contain the percentage of detected signals for each angle at which the phone was held.

	0 (downwards)	45	90	135	180 (upwards)
1	100	100	100	100	100
2	100	100	100	100	100
3	80	100	100	80	100
4	90	100	100	90	100
5	100	90	100	100	100
6	100	100	100	100	100
7	90	100	100	70	100
8	100	100	100	100	100
9	70	100	100	100	100
10	30	80	100	100	100



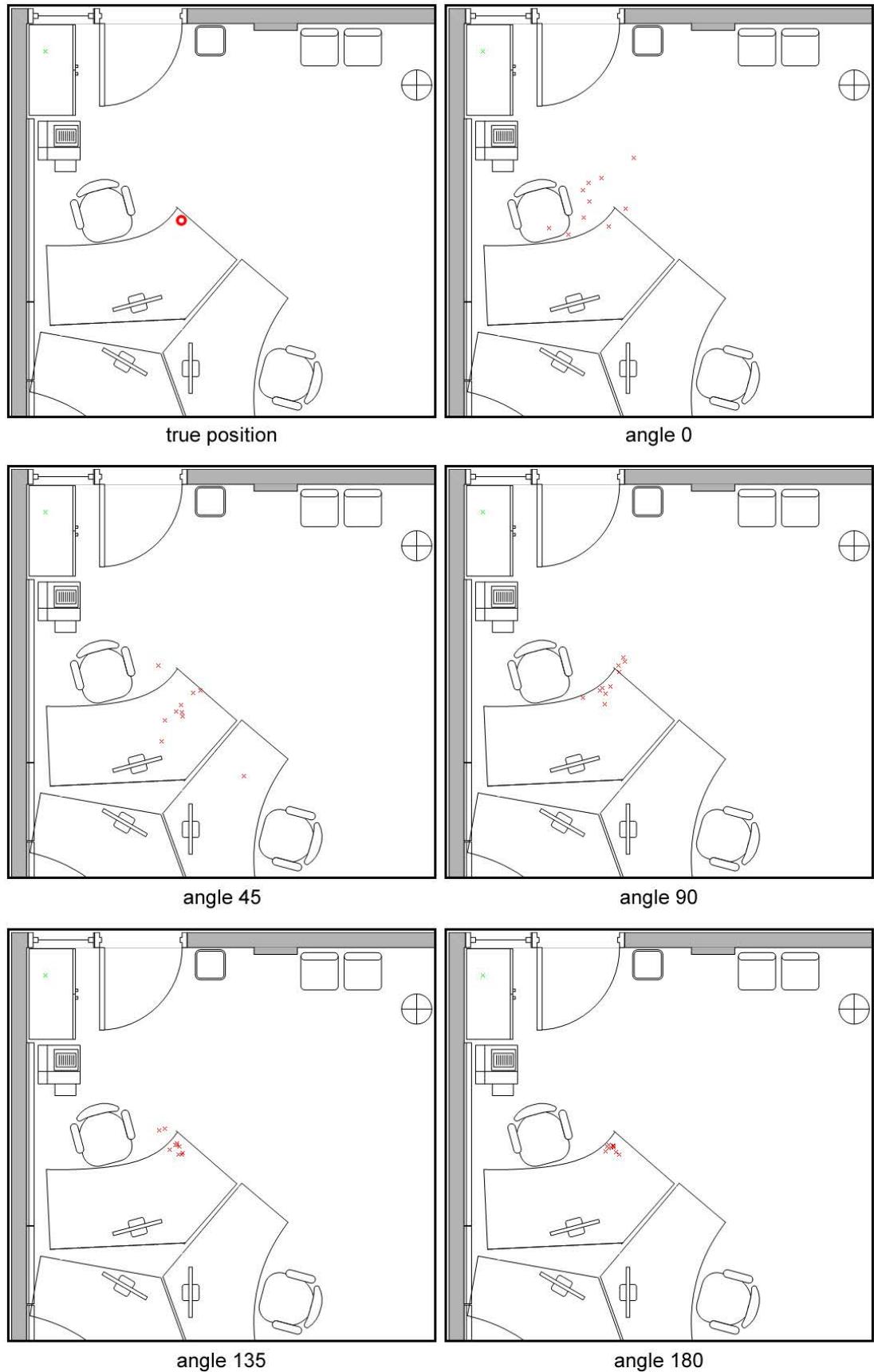
**Figure 39: The change of distance (in cm) between MPV and correct position as the speaker is rotated from downward to upward orientation. Each colour represents one control point. Vertical axis is distance and is measured in centimetres. Horizontal axis is angle of rotation with zero being downward and 180 being upward. Notice how accuracy increases as speaker angle approaches 180 degrees (upwards). In all cases positioning accuracy remains sub metre.**

Angles 90 (horizontal) and 180 (upward) are the only angles that have 100% detection. However 90, unlike 180, has a huge spread in accuracy from some of the worst results to results on par with upward orientation. Interestingly accuracies for angle 90 tend to be either very good or very bad, which illustrates the importance of direction when dealing with sound waves. The worst accuracy for 90 degree angle was recorded at point 5, which is roughly the centre of the room. Angle 0 (downward) had the worst detection rate. Accuracies for some points at angle 0 may appear misleading. For example point 10, which is the farthest point from the centre of the room appears to have good accuracy (24 cm) at angle 0, however detection rate for this point and angle is only 30%. Angle 45 has a much better detection rate than angle 0 and the spread in accuracies is the greatest compared to every other orientation. However unlike angle 90,

it is very evenly distributed. Angle 135 has a worse detection rate than angle 45, its distribution of accuracies is also very even, however the upper boundary is less than half of the upper boundary for angle 45. From a practical point of view, this angle has little importance, because the user is unlikely to hold the phone in such a way unless instructed to do so.

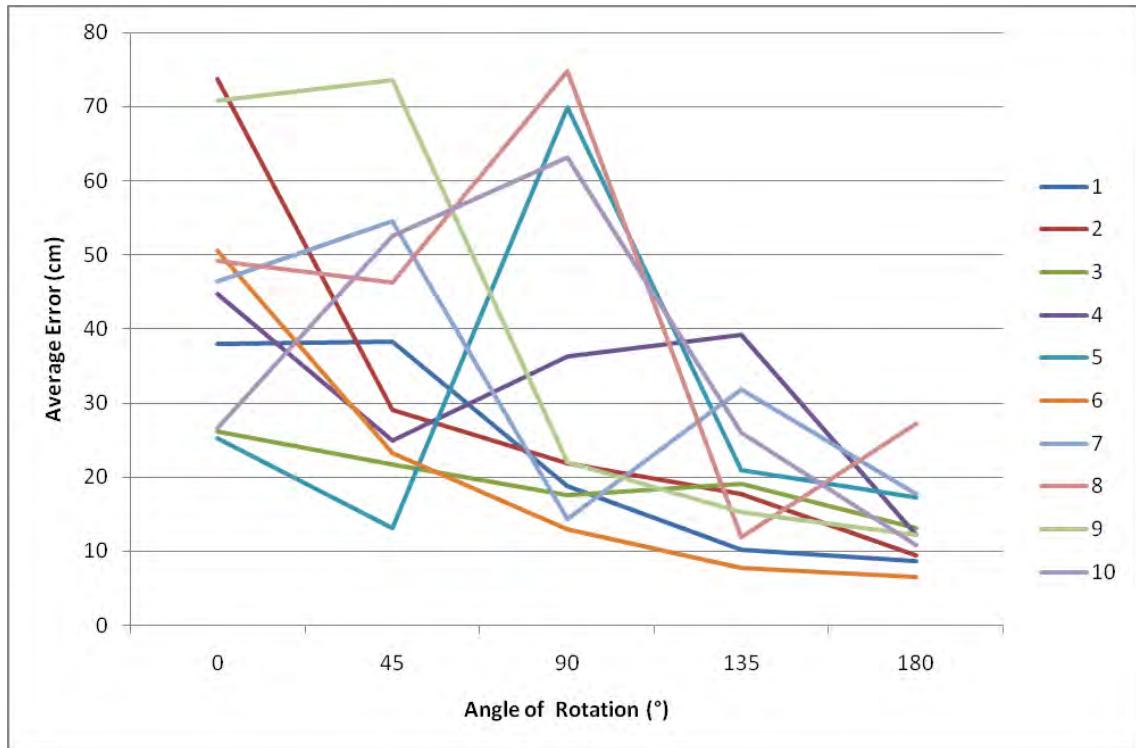
In terms of accuracy, detection rate, and consistency, angle 180 (upward) is undoubtedly the most reliable orientation. This makes us think that unless some parameters of the system are changed such as increased sound volume, less directional microphones or lower signal detection thresholds, the most accurate approach would be to instruct users to momentarily flip their phone upside down and employ the phone's accelerometers to detect this flip to trigger the positioning procedure. Although this may not be very good practice in terms of user interface interaction, the increased accuracy and reliability is a valuable trade off for enhancing the user's navigation experience. However, it should be noted that overall positioning accuracy is still sub metre in all speaker orientations.

Unlike the previous test where detected positions were clustered very closely together, some combinations of angle and position, in particular angles below 90 and positions farther away from the centre of the room, produced scattered results. For example, the scattering for point 1 can be seen on Figure 40. This means that distance to MPV alone does not very well reflect actual accuracy and reliability. For example 10 detected positions could be scattered across the room, but MPV based on these 10 positions could by chance be located very closely to the correct position. Therefore an average of the distances between the correct position and each of the 10 detected positions would better reflect the accuracy of the test. See Figure 41.



**Figure 40: The scattering of detected positions and various pitch angles.** The red circle indicates the true position. Red crosses mark each individual position calculation.

It appears that the biggest difference between the two charts is the result for point 6 at angle 45. On the chart with distance to MPV the result was extremely accurate, however on the chart with average distance to each detected position it is a lot less accurate. Point 8 at 90 degrees also appears to be a lot less accurate. Otherwise the two charts look very similar.



**Figure 41: The change in the average of distances (in cm) between each detected position and correct position as the speaker is rotated from downward to upward orientation. Each colour represents one control point. Vertical axis is distance and is measured in centimetres. Horizontal axis is angle of rotation with zero being downward and 180 being upward.**

Either method of accuracy estimation doesn't give errors above 80 cm. For upward orientation the error doesn't exceed 30 cm. Or 20 cm, if point 8 is considered an outlier.

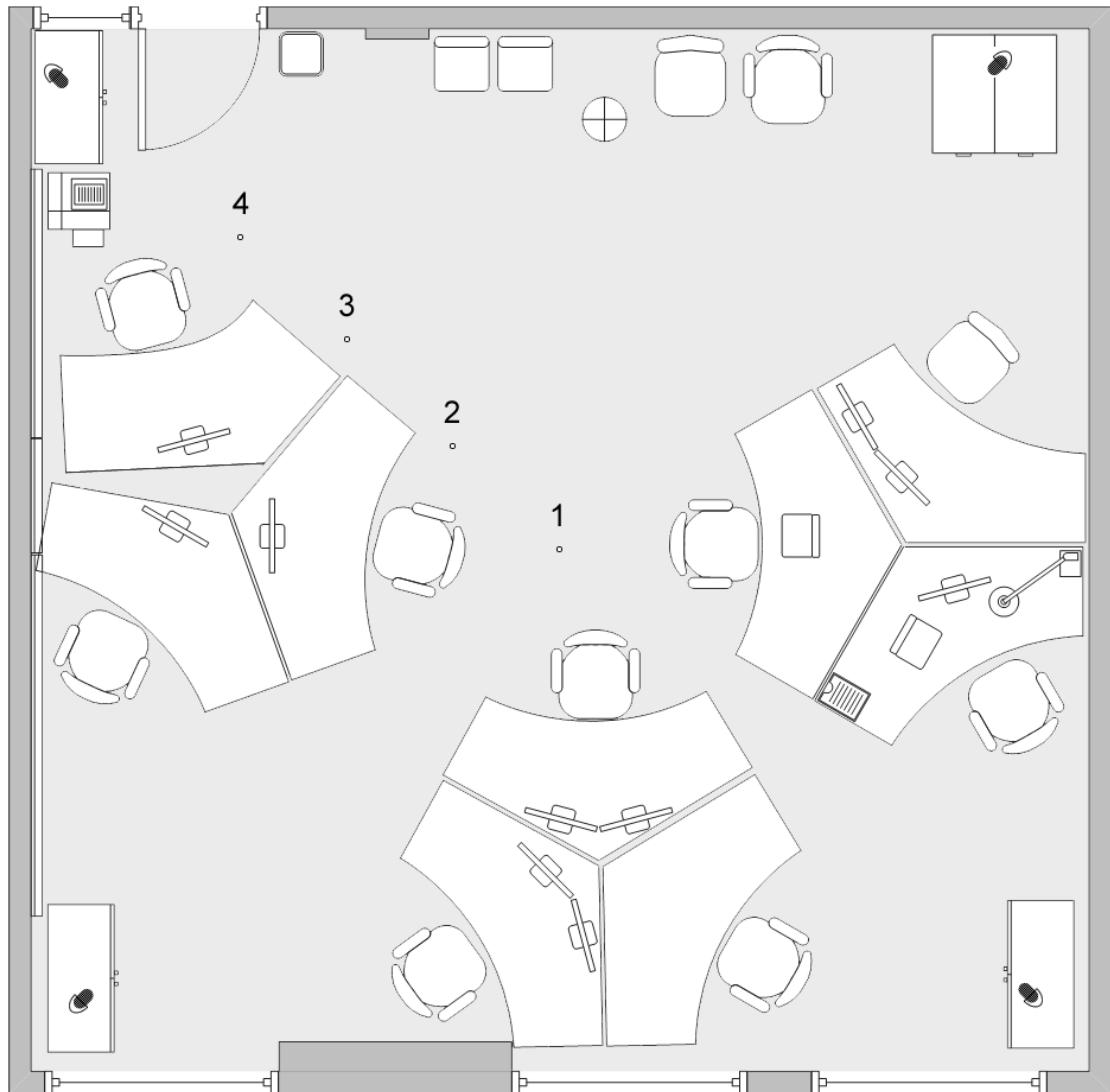
### 5.3. Direction and Signal Obstruction

This section explores two factors that were excluded from experiments described in the previous section so as to reduce the number of variables:

1. **Yaw orientation of the phone.** Ultrasound is very directional and orientation of the speaker in relation to a microphone has significant effect on reception quality. Unless the speaker is pointing directly upward or downward, its yaw orientation is going to impact reception quality at each of the four microphones. This in turn can influence quality of signal arrival timestamping and therefore trilateration accuracy.
2. **User obstructing direct line-of-sight between the phone and one of the microphones.** Unlike low frequency sounds that have a long wavelength, ultrasound has poor obstacle penetration. This means that maintaining a direct line-of-sight between the phone and all four microphones is important in order to achieve best accuracy. Blocking direct line-of-sight to one of the microphones can negatively impact signal intensity on reception and introduce lag due to signal arriving via diffraction. Both poor reception and lag can reduce trilateration accuracy. The most likely scenario where one of the microphones would be blocked is when the user himself is standing in the path of the signal. Although microphones were placed on the ceiling to address this problem, in some cases direct line-of-sight can still be blocked, for example if the user is standing too far from a microphone.

**Experiment Method.** Positioning was done in the same lab as in the previous section. However a different set of check points was used this time. First check point (P1) is located precisely in the centre of the room. Each next check point is one meter closer to

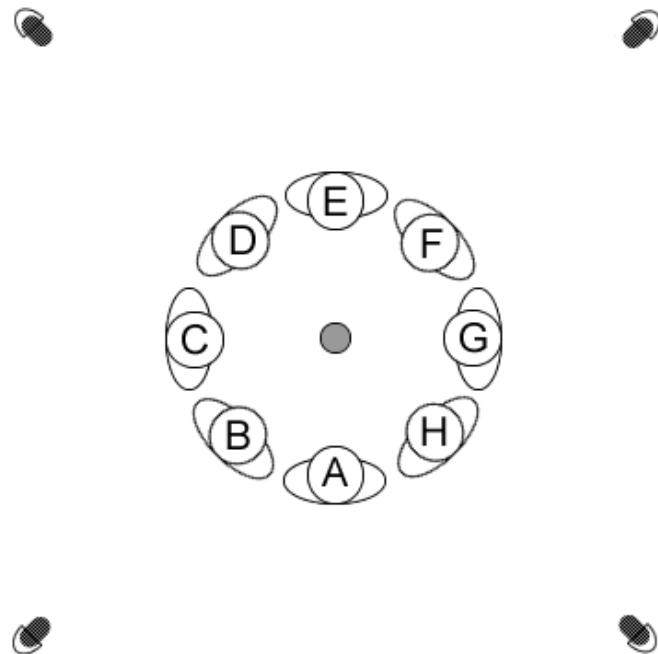
one of the corners. There are altogether four check points (P1, P2, P3 and P4). See Figure 42.



**Figure 42: Layout of the room used for direction and signal obstruction experiments.** Dimensions of the room are 7.1 by 7 meters. The position of 4 check points is shown as well as the position of the four microphones placed at near ceiling height in each corner.

At each of the four check points the user stood in 8 different ways while holding the phone directly over the centre of the check point (Figure 43). In positions B, D, F, H the user would stand in the direct line-of-sight between the phone he is holding and one of the microphones. These positions are later referenced to as “blocked”. The actual path of the signal would either be blocked by the user’s body or pass above his head

depending on how far he is from the given microphone. In positions A, C, E, G the user stands in such a way as to not block the path to any microphones. These positions are later referenced to as “direct”. An example of positions for check point one can be seen in Figure 43. For other check point positions the user’s position was adjusted so that in blocked positions the user stands directly between the phone and a microphone and in direct positions he stands exactly between two speaker-microphone lines of sight. Because the shape of the room is very close to a square, positions D and H remained in the same place in relation to each check point.



**Figure 43: Positions in which the user stands at check point 1.** Grey circle in the centre shows where the smartphone is held. The 8 silhouettes marked A-H show where the user stands with their backs facing outwards in all cases.

In each of the 8 positions the phone was held in 3 different orientations: with the speaker facing upwards (up), facing the user (back) and facing away from the user (forward). In each orientation the ultrasound signal was produced 10 times, which are treated as one sample. Altogether in the experiment 960 signals were sent, making 96 samples.

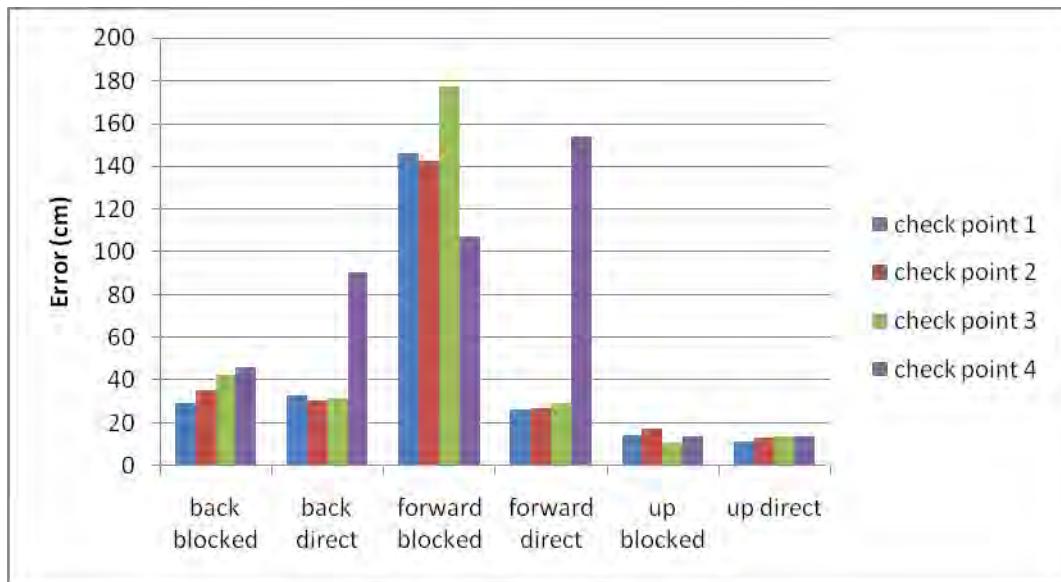
**Discussion.** For each sample the Most Probable Value (MPV) was calculated. After this the difference between MPV and true position of the check point was calculated. The results for each checkpoint were divided into two groups: blocked (B, D, F, H), where the user stands in the way of one of the microphones and direct (A, C, E, G), where the user outside the direct path. The results can be found in Table 8 and Figure 44. The percentage of failed positioning attempts was also calculated for each of the 96 samples. Their comparison can be found in Figure 45.

**Table 8: Difference between MPV and true position  $\pm$  standard deviation for each check point, user's position category and orientation.** First column identifies combination of phone orientation and the type of user's position. Columns 2, 3, 4, 5 contain results for each of the four check points in meters.

orientation	check point 1	check point 2	check point 3	check point 4
back blocked	$0.293 \pm 0.109$	$0.353 \pm 0.115$	$0.424 \pm 0.071$	$0.460 \pm 0.075$
back direct	$0.330 \pm 0.085$	$0.305 \pm 0.082$	$0.318 \pm 0.081$	$0.907 \pm 0.232$
forward blocked	$1.464 \pm 0.149$	$1.428 \pm 0.131$	$1.771 \pm 0.160$	$1.074 \pm 0.103$
forward direct	$0.266 \pm 0.080$	$0.270 \pm 0.108$	$0.295 \pm 0.044$	$1.539 \pm 0.222$
up blocked	$0.142 \pm 0.020$	$0.172 \pm 0.047$	$0.110 \pm 0.025$	$0.141 \pm 0.041$
up direct	$0.114 \pm 0.015$	$0.135 \pm 0.030$	$0.139 \pm 0.033$	$0.141 \pm 0.016$

In terms of accuracy, upward speaker orientation gives consistent results in the 10-20 cm accuracy range regardless of where the user stands, which suggests that in this orientation, a user's position has little or no effect on accuracy. Regarding the other four configurations, forward blocked is consistently very inaccurate across all four check points. Most probably this is a direct result of a very disadvantageous combination of the user's position and speaker orientation. Ultrasound is very directional and it will have the least energy directly behind the speaker. In forward orientation, the user stands directly behind the speaker, making the signal even weaker in this direction. Finally, blocked position means that there is a microphone directly behind the user. In back blocked configuration a microphone that is directly in front of the speaker is the one being blocked, which makes the two factors cancel each other out. As a result back

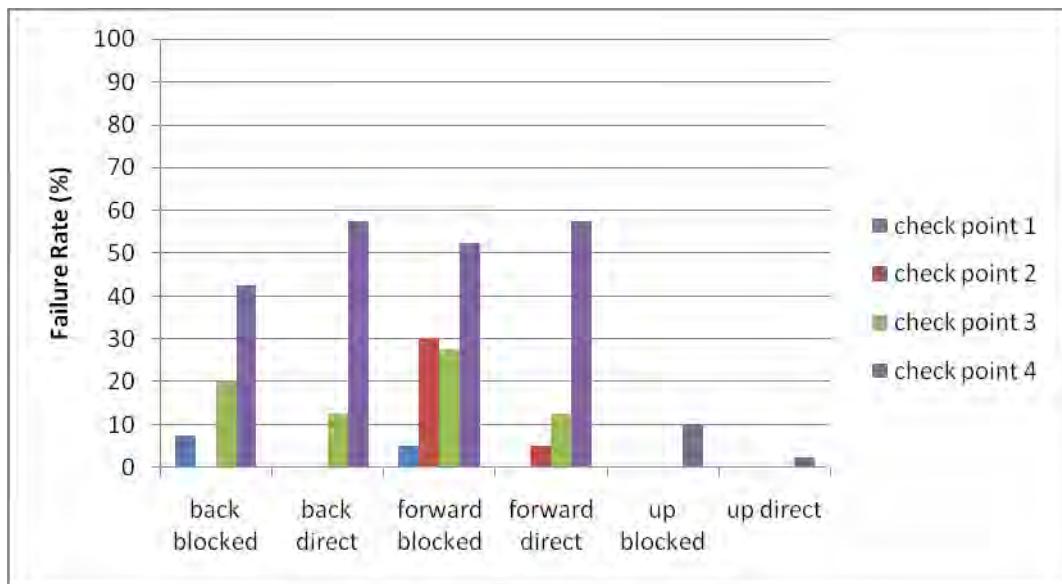
blocked and the other two horizontal configurations where the user doesn't block any microphones show very similar accuracy for check points 1-3 ranging between 20 and 40 cm. Results for check point 4 show worse accuracy ranging from 40 cm to 1.5 meters, however these results are less reliable since around half of the signals were not detected which resulted in a smaller sample size.



**Figure 44: Difference between MPV and true position for each check point, user's position type and orientation.** Each colour represents a check point. Vertical axis represents error in centimetres. Horizontal axis contains all combinations of phone orientation (facing the user, facing away from the user and facing the ceiling) and types of user position (blocking or not blocking a microphone).

Analysis of failure rate reveal similar trends with the discrepancy between the first three check points and the last check point being more pronounced. Upward orientation has 0% failure rate at check points 1-3. At check point 4 there was 2.5% failure rate in direct configuration and 10% failure rate in blocked configuration. Because the phone is far from the centre of the room at check point 4, signal reception is very weak at the opposite corner and almost just as weak at the two other corners. As the distance to the three microphones increases, the direct line-of-sight gets a more gradual slope and the user can obstruct it more, which explains the higher failure rate in blocked positions. Similarly to accuracy results, the forward blocked configuration has the worst results.

Failure rate of other non-upward configurations didn't exceed 20% for check points 1-3, with check points 1-2 only going as high as 7.5%. Failure rate at check point 4 for all non-upward configurations was very high ranging between 40 and 60 percent. Because the results are equally bad for both blocked and direct positions, it can be concluded that high failure rate was primarily the result of most of the signal's energy being directed away from the three distant microphones. This is very likely to happen in any orientation other than upwards if the user is not paying attention to where the speaker is pointing. Although upward configuration does not give the most efficient distribution of energy, it guarantees certain minimum levels at all four microphones which is necessary for ultrasonic positioning to work reliably.



**Figure 45: Percentage of failed positioning attempts for each check point, user's position category and orientation.** Each colour represents a check point. Vertical axis represents percentage of failed positioning attempts. Horizontal axis contains all combinations of phone orientation (facing the user, facing away from the user and facing the ceiling) and types of user position (blocking or not blocking a microphone).

Poor accuracy and high failure rate should not be viewed as two separate problems. Whether positioning fails or performs inaccurately depends on implementation. Either of them indicates that conditions for trilateration were unfavourable in a given situation.

Given that the user will not pay attention to where he stands in relation to the microphones the following observations can be made:

- Speaker pointing upward is the most accurate and reliable orientation.
- The worst case scenario is the user holding a phone in such a way that the speaker faces directly away from him.
- Accuracy and reliability deteriorates as the user moves further from the centre of the room.

#### **5.4. Background Noise**

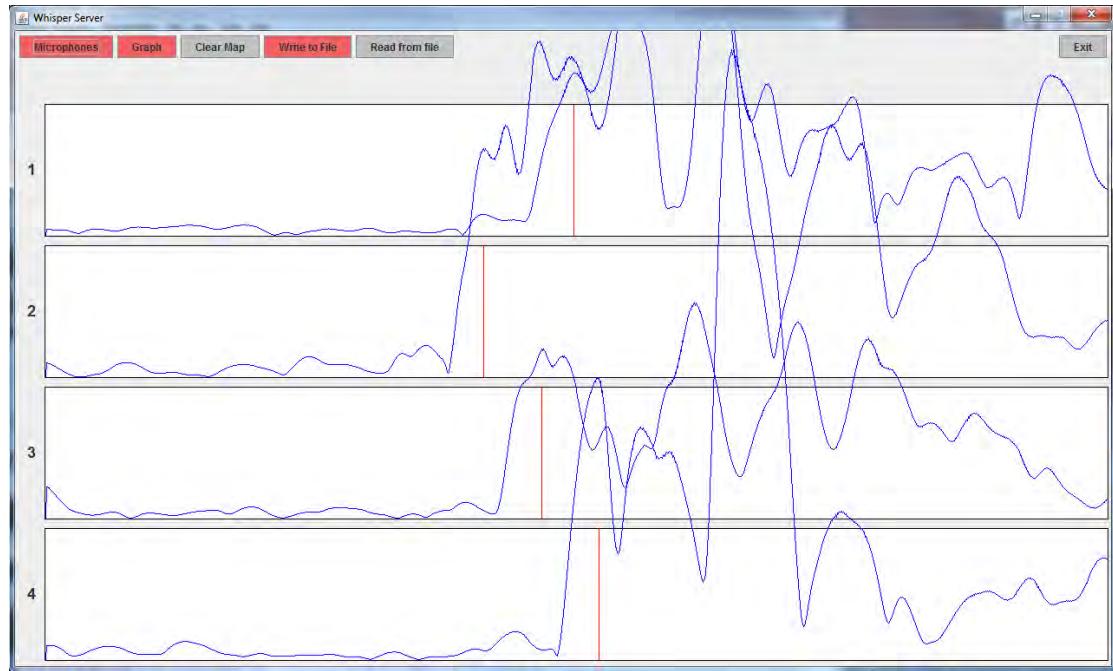
This section addresses the seventh research question (RQ7): Can background noise cause false positives and how can this be countered? Positioning systems that use sound are vulnerable to loud background noise. Current implementation of the positioning system reads four audio streams into buffers, performs a bandpass filter and looks for a spike in the given frequency in order to detect signal arrival times. Therefore if the utilised frequency accidentally appeared in background noise, it could be erroneously confused with the real signal. Fortunately sounds that regularly occur in office environment such as noise from working computers, ventilation, talking, walking and typing have not been observed interfering with the system. Only the following kinds of noise were observed triggering signal detection in the program:

- Clanking noises resulting from small metal objects hitting each other generated a very strong response. For example jiggling a bunch of keys (see Figure 46). It appears that these sounds have a mostly ultrasonic nature and strongly overlap with the chosen positioning frequency.

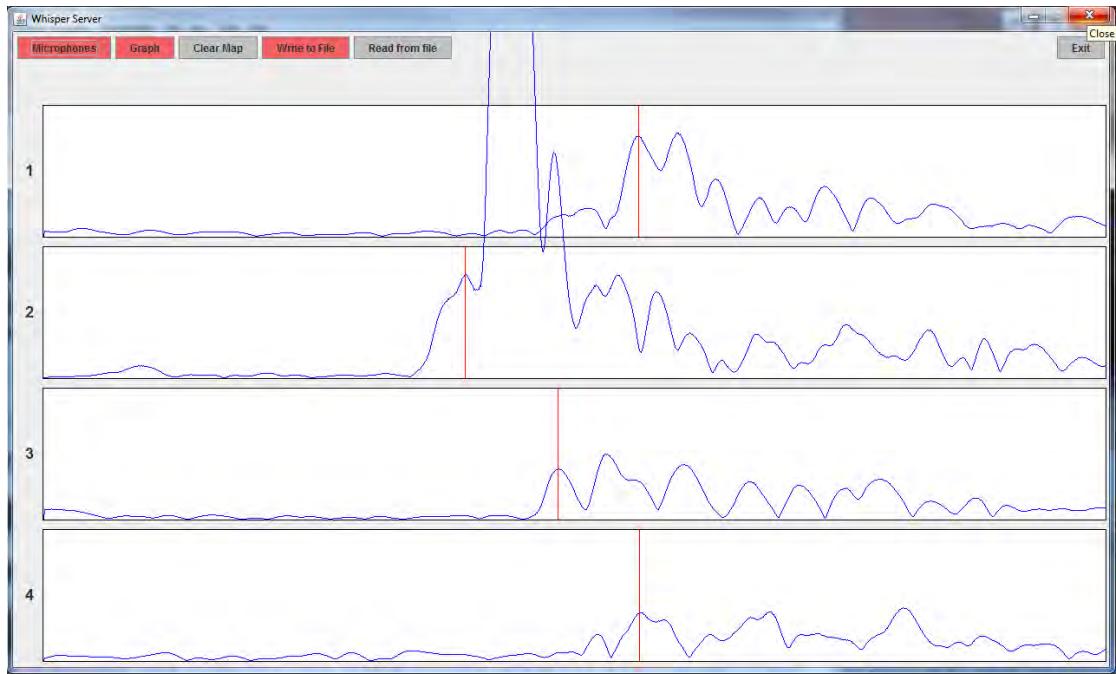
- Loud noises resulting from slamming shut a drawer or a door (see Figure 47).

These sounds have a broadband nature and are very loud.

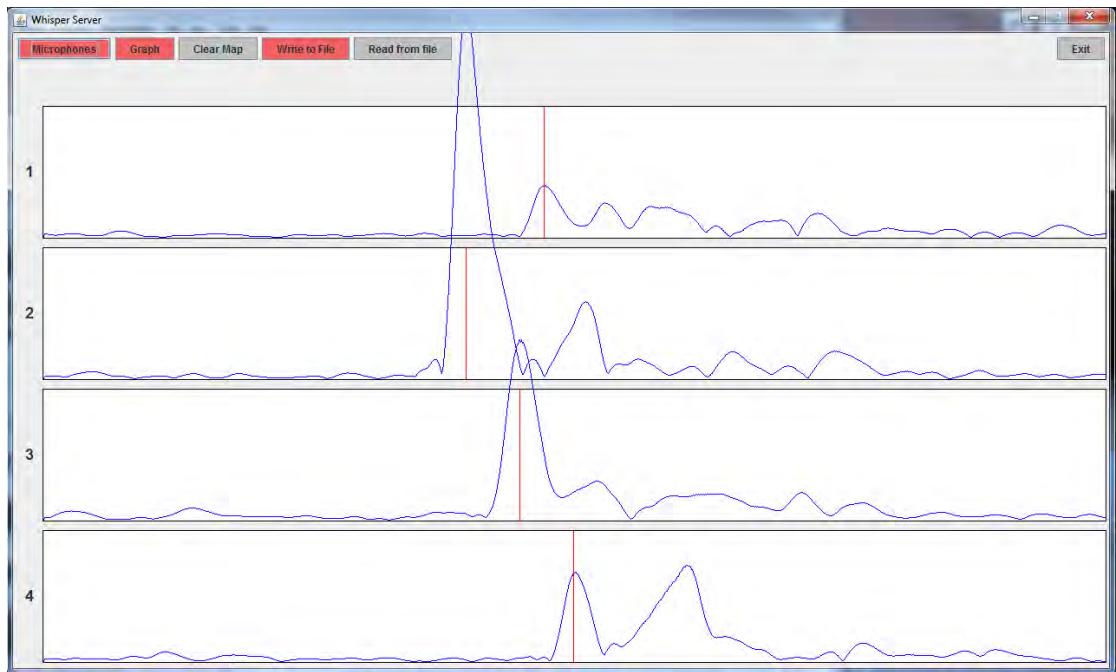
For comparison with regular signal detection see Figure 48.



**Figure 46:** Noise generated from jiggling a bunch of keys after 21.5 kHz bandpass filter. Each graph frame represents one of the four channels. The blue line shows intensity of the chosen frequency in the current time frame. Vertical red line shows estimated point of signal's arrival.



**Figure 47:** Noise generated from slamming shut a metal drawer after 21.5 kHz bandpass filter.



**Figure 48:** Signal detection in normal room noise conditions.

Wi-Fi communication between mobile phone and the positioning system was not used during most experiments. There was no way for the system to know when the next ultrasound signal will be sent and the signal detection process was running constantly.

This made it very vulnerable to false detections resulting from interference with noises listed above. In a practical scenario where the user expects the system to relay his position to the phone, false detection can be almost completely avoided on the client side by sending the latest detected position in response to a request. A smartphone would first send an ultrasound signal, wait for a certain period of time and send a request for his position over Wi-Fi. If the delay was calibrated correctly position sent back to the phone will be the correct position most of the time. It will be incorrect only in one of the three cases:

- Ultrasound signal was not received. This can happen if the delay was too short, there was an obstacle in the signal's path or the speaker malfunctioned.
- An interfering noise caused a false detection in the short timeframe between ultrasound signal arrival and position request.
- An ultrasound signal from another phone was detected in the short timeframe between arrival of the correct ultrasound signal and position request.

There are alternative precautions that can be taken in order to minimise impact from interfering noises.

**Limit detection timeframe.** False detection can be reduced by switching detection on only when necessary. For example the phone would communicate its intent to send the ultrasound signal, the server would switch detection on and reply that it is ready. Once the signal was received at all four channels, detection would be switched off. By making detection only work for short periods of time and only when necessary the possibility of encountering interference from random background noises can be reduced

dramatically. This method is not necessarily better than the method described above as it is going to fail under the same conditions.

**Use pattern recognition.** In scenarios where RF communication between client and server is not possible it becomes very hard to eliminate false detection of background noise. Virtually any noise that happens to contain the frequency used by the signal will be confused with the real signal. Making the signal a combination of spikes in several frequencies divided by delays of a certain length and looking for this exact pattern during signal detection is potentially an effective method to eliminate false detection, because it is very unlikely that this exact same pattern will occur naturally as a part of background noise. This method is also considered as a potential solution for supporting a large number of users as a unique pattern can be assigned to each phone in the room (See Section 6.3.2).

## 6. CONCLUSIONS

### 6.1. Summary of Work

In the beginning of this research, designing, developing and testing a novel accurate indoor positioning approach was identified to be the primary goal. Commercial off-the-shelf (COTS) mobile phones with no hardware or Operating System level modifications were chosen as the desired computing platform. This would allow anyone who carries a mobile phone with them to be in possession of the required hardware. Once the software component is installed they can participate. Another advantage is that a mobile phone, in particular a smartphone, is a powerful interface for an LBS that can exploit the positioning system, something many contemporary Indoor Positioning Systems lack.

Because of the choices mentioned above the list of potential technologies was limited to only those that are present on a majority of modern smartphones: Satellite Navigation Systems, GSM, Wi-Fi, Bluetooth, Sound, Dead Reckoning and Computer Vision. Each technology was researched and reviewed separately in the literature review. It was observed that so far no indoor positioning system has been able to locate a regular COTS phone, that doesn't include any rare hardware upgrades such as WiMax, ZigBee or Bluetooth 4.0, with positioning accuracy below one meter (sub-meter). It was also noted in the literature review that ultrasound trilateration could potentially pass this 1 meter barrier with ease, which to the best of our knowledge and from the literature has not been attempted with mobile phones anywhere before.

Before development of the positioning system could start, a number of factors had to be investigated through experiments:

- It was unknown how well mobile phones would be able to produce ultrasound. The ability to produce ultrasound frequencies is not officially supported by manufacturers. It could be viewed as a byproduct of the way sound hardware works. Four different mobile phone models were tested at a number of volume settings. It was observed that all four phones were able to produce ultrasound signals in the 17-22 kHz range. However under maximum volume settings some audible noise was generated along with the signal. Some phones produced a lot of noise and for some the noise was very mild, however reductions in volume eliminated audible noise for all four phones without exception. Altogether the results were positive as it was found possible to produce an inaudible audio signal on a COTS mobile phone and detect it with a microphone.
- There are a number of signal properties such as shape, length, and frequency that can influence how easy or hard it is to detect and correctly timestamp the arrival of an ultrasound signal. Also these properties determine how likely the signal is to be audible by users. These factors were taken into account when designing the ultrasound signal that was used in all later stages. Making the signal envelope use a different frequency than the reference part of the signal substantially improved timestamping accuracy.
- High frequencies are known to attenuate very fast when travelling through air. This is particularly problematic because mobile phones have very limited volume levels. Also high frequencies are very directional which means that anywhere, except right in front of the speaker, sound levels are weaker. It was therefore necessary to find out at what distance the signal can be reliably detected regardless of the speaker's orientation to the microphone. This determined the maximum diagonal length of an area that could be covered by the positioning system. Fortunately it was possible to

cover the entire lab, which is 7 by 7 meters. For larger rooms than this, more microphones could be installed to ensure complete coverage.

After the preliminary experiments were completed and it was concluded that a phone's position can be found anywhere in the lab using only four microphones, development of an ultrasound positioning system started. The first step was to establish a way for our program to access a sound card and stream data from four channels simultaneously. This data would be stored in a buffer and all manipulations would have to be done in the timeframe between the last and the next update in order for the system to work in real time. The next step was signal detection. This was done with a Bessel bandpass filter. A signal would be considered detected if the filter output reached a certain threshold after which some checks would be done to separate the reference point from effects of multipath. The threshold was tweaked in order to find the best compromise between reliable signal detection and false detections caused by various noises such as a door being slammed.

Being able to determine the point at which the signal arrived at each of the microphones meant it was possible to calculate the difference in distances from the phone to each microphone. In order to calculate position using only the difference in delay and the precise position of each microphone a novel Time Difference of Arrival (TDoA) trilateration algorithm was designed. This made it possible to avoid synchronisation with the phone, which would have introduced inaccuracy and complexity to the positioning system. The trilateration method was first implemented and tested as a standalone program as a proof of concept and later integrated into the positioning system. After this last step was taken, the server side of the system was able to determine the phone's position in two dimensions in a real-world environment.

During most of the client side development of the system, a very simple program would play a “WAV” file with the signal either once or in a loop. Later a proper client was developed that would detect the phone being flipped (or the screen being tapped), produce the signal, send a request to the server over Wi-Fi, receive the position in the form of coordinates and display it on the phone’s screen as a red dot overlaid on the plan of the room in real-time.

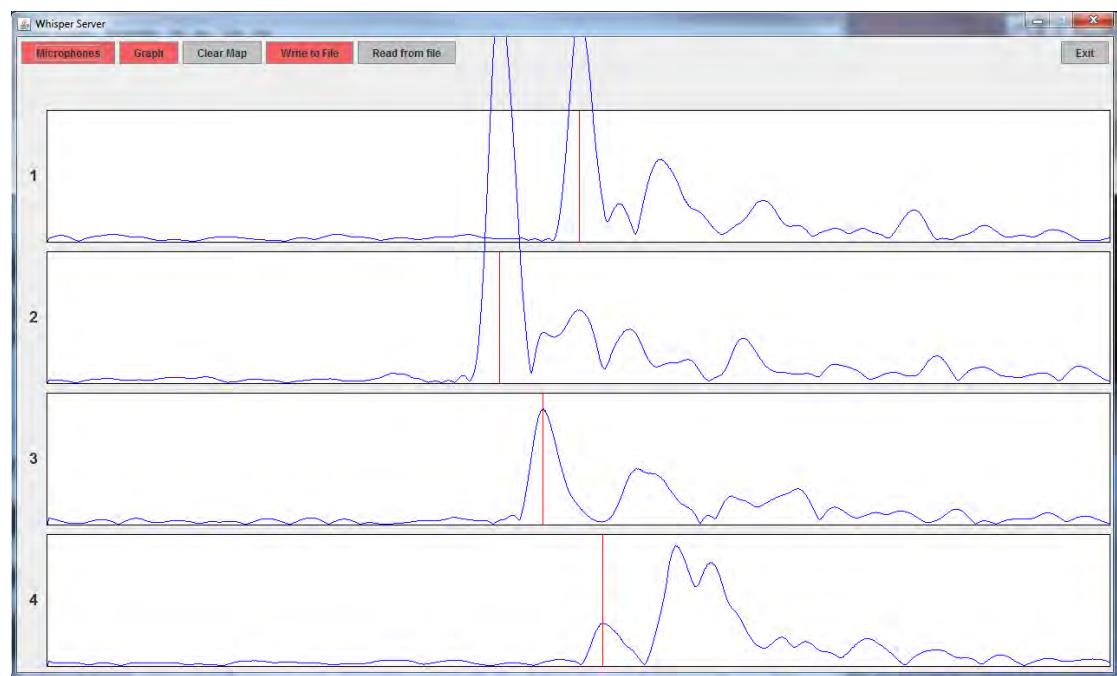
The final stage of this research involved testing the positioning system in order to determine its accuracy and reliability as well as to find out merits and shortcomings of mobile phone ultrasound positioning in general and of this implementation in particular.

The following tools were available for this purpose:

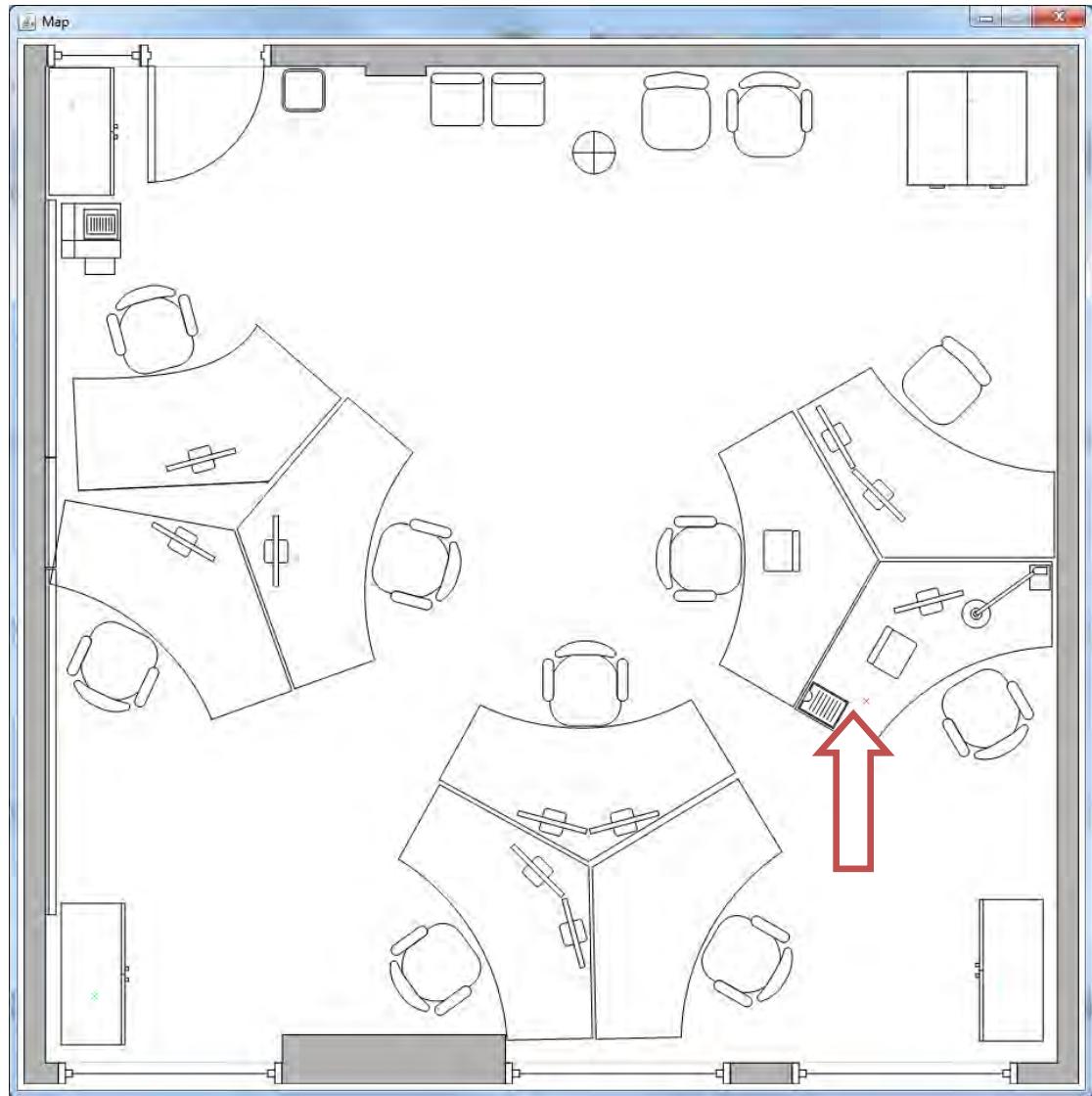
- A real-time graphical representation of the four buffers after the bandpass filter. It can be frozen at any moment to facilitate thorough analysis. This feature is useful for low-level troubleshooting and to monitor noise levels, signal strength and delay. See Figure 49.
- A plan of the room that is automatically updated with the latest position fix. The position is shown with a red cross. This tool can be used to collect multiple readings taken at a single location and visually analyse their spread and accuracy or monitor positioning output in real time. See Figure 50.
- A function that dumps all program output into a text file. Information recorded consists of estimated x and y coordinates as well as standard deviation. This data can be imported into an Excel spreadsheet. In that form it can be thoroughly analysed in order to find relevant trends or patterns.

Pictures of the lab and equipment can be found in Appendix 5.

The positioning system was tested for accuracy with three different settings at 10 different known positions in the room. For each point an average, best and worst result were calculated. These results were used to analyze and compare the three settings. In a different experiment the phone was held at various angles in order to identify the phone orientation that gives the best positioning reliability and to evaluate disadvantages resulting from other orientations.



**Figure 49:** Screenshot of a program window that displays current buffer contents after filtering. Each graph frame represents one of the four channels. The blue line shows intensity of the chosen frequency in the current time frame. Vertical red line shows estimated point of signal's arrival.



**Figure 50:** Screenshot of a program window that displays detected user position. Plan of the room where positioning experiments take place is displayed in the background. Red crosses show detected user position. Green crosses show the location of microphones.

## 6.2. Contributions of the Thesis

This research makes two novel contributions to the field of indoor positioning.

### Ultrasound indoor positioning for mobile phones.

A proof-of-concept positioning prototype was developed and tested during the course of this research. It was demonstrated that an off-the-shelf mobile phone can be located in a

7 by 7 meter room with better than 10 cm accuracy using four microphones, a sound card and an average PC. The ultrasound signal used in the locationing process cannot be heard by anyone in the room as its frequency is above the range normally audible to humans. At the same time the frequency is low enough for regular audio hardware to reproduce and detect it. This means that theoretically any mobile phone is compatible with this positioning method unless reproducible frequency range was limited by the manufacturer for some reason. To the best of our knowledge no other indoor positioning system can locate a regular smartphone that doesn't include any rare/experimental hardware, with comparable accuracy. Given that the described system is only a prototype, the method can be further developed and implemented as a network of wirelessly synchronised beacons each carrying a microphone. Anybody who enters the covered area with a smartphone will be able to immediately take advantage of the positioning system and any LBS it enables by simply downloading and installing the application.

### **Asynchronous trilateration.**

An asynchronous trilateration algorithm was developed that allows for locationing of a signal source in two dimensions using time-of-flight without the need to know the time the signal was sent. The combination of this algorithm applied to indoor positioning on COTS mobile phones using ultrasound is not found in the literature, which makes our approach a contribution to the state-of-the-art in this research field. Accuracy was shown to be comparable to standard least squares trilateration and the ability to avoid synchronisation between signal source/receiver coming at the price of one extra control point (microphone). Asynchronous trilateration was directly derived from synchronous least squares trilateration and introduces no additional complexity or prerequisites other than those required to avoid synchronisation. This makes it advantageous over existing

methods such as Bancroft which was developed as a solution for a specific problem (Bancroft 1985).

### 6.3. Discussion of Results

The developed positioning system prototype can be evaluated in terms of the use case given towards the end of section 1.2. All four scenarios are forms of indoor LBS and as such take advantage of mobile phone's ability to locate itself in a building. Scenarios 1 and 2 involve pointing the phone at an object and doing directional querying. Scenarios 2, 3 and 4 involve guiding a user around the building.

As a source of coordinates for directional querying our approach has advantages and disadvantages. Centimetre level accuracy means that objects the size of bottle, vase, lamp can be correctly queried, provided that orientation of the phone can be determined with great accuracy. Also a positioning fix can be done almost instantly without the need for the system to run in the background, which is ideal for scenarios like directional querying. The fact that the phone should be held upside-down to get the best accuracy is a disadvantage. Doing a positional fix with the screen up, which is how most users will do directional querying results in accuracy dropping to 30-40 cm. This is still very good accuracy, unfortunately this orientation may result in reduced reliability in the form of failed positional fixes.

Navigating in a building requires continuous tracking of the phone. With the current implementation it is possible to do a positional fix twice a second without the signal arriving via multipath affecting the next positional fix. Such refresh rate is deemed sufficient for most navigational tasks. It has not been researched how well our

positioning prototype tracks moving targets because there is a lot of room for optimisation e.g. take into account the user's previous position, calculate the probability of a positional fix being correct based on user's trajectory, increases the number of fixes per second etc.

Although ultrasound positioning in most cases won't be able to work from inside a bag or pocket, this is not a problem in any of the given scenarios. Because tracking can be initiated and stopped momentarily there is no need for the system to be able to run in background. As soon as the user takes out the phone and starts to interact with the app, his position will be immediately calculated and used by the program. The ability to work in the background is only useful if we want to track and record how people move in the building which was not among our objectives.

## 6.4. Future Work

Several directions were identified in which research presented in this thesis can be continued.

### 6.4.1. Directional querying

Location is not the only type of spatial data that can be useful in a Location Based Service. Orientation together with position can be used to calculate direction, which among other things can be used for directional querying, a powerful LBS application. Most smartphones carry magnetometers and accelerometers which together can be used to calculate orientation of the phone. Pitch and Roll can be easily detected using inbuilt accelerometers. Although most of the time accelerometers are used to detect changes

between portrait and landscape orientation of the device, they are also successfully used in games to detect Pitch and Roll simultaneously with great precision. Yaw is a lot harder to calculate correctly, because unlike Pitch and Roll there is no strong omnipresent reference such as gravity. An equivalent of gravity for Yaw would be the Earth's magnetic field, which is very weak and is easily distorted by large metal objects, electric devices and magnets. This magnetic field is used by magnetometers to detect the direction to magnetic North or in other words Yaw. Some smartphones also come with gyroscopes which when activated can very accurately track rotation of the phone in 3 axis.

Magnetometer or gyroscope alone cannot be used to accurately determine Yaw of the phone at any given time. Magnetometers are easily distracted by local magnetic fields which are abundant in indoor environments. Gyroscopes don't have a reference point. Although they can accurately track the phone's orientation in relation to their orientation at the moment of activation, this initial orientation is unknown. A reference would have to be set every time the gyroscope is activated.

We propose a method for accurately tracking Yaw orientation of the phone using a combination of magnetometer, gyroscope and the indoor positioning system outlined in this thesis. Magnetic anomalies tend to stay the same in indoor environments unless large furniture and equipment is moved<sup>31</sup>. This means that if a magnetometer is affected by a local magnetic field, the direction to magnetic north will be determined incorrectly,

---

<sup>31</sup> "IndoorAtlas. Ambient magnetic field-based indoor location technology" Retrieved 12 November, 2012, from <http://web.indooratlas.com/web/WhitePaper.pdf>

but the number of degrees by which it is offset will remain constant in the same location. Provided that the location of the phone is known, which can be done using our indoor positioning system, and the angle offset, which can be measured beforehand; correct Yaw can be determined on a smartphone. Thanks to gyroscopes it may not be necessary to know the correct offset for every location in the room. The offset may be measured for only several key locations, such as the doorway or any other point a user is guaranteed to walk through. When the positioning software is activated, it will keep checking if the user is at one of the reference points. When such an event is detected, it will take measurements from the magnetometer and using the known offset calculate the true direction to magnetic north. This direction could subsequently be used with the readings from the gyroscope to determine Yaw. The direction will be updated if necessary, every time a user passes through one of the reference points in order to eliminate accumulated gyroscope drift.

Access to both accurate position and orientation will make directional querying applications a possibility. For example if a user was in a museum, he could point his phone at an exhibit and click a button in order to find more about it. The positioning system will detect the phone's position, orientation and draw a virtual line in the direction it is pointing. Provided that the exhibit is registered with the spatial database, the line will intersect with the exhibit's bounding box and querying will be successful. The user will then receive multimedia content relevant to the exhibit in the form of images, video, audio, text, hypertext, links to relevant Wikipedia articles, etc.

### 6.4.2. Support Multiple Users

Currently the positioning system uses identical ultrasound signals and therefore is unable to determine which phone the signal came from other than by assuming that a position request sent over Wi-Fi was produced by the same phone as the most recently detected signal. Although it is not very likely that another phone will produce the signal during the gap between the ultrasound being produced and the request being received, the design requires some improvement in order to support multiple users. Two different approaches were identified.

**Queue.** The currently used method can be made more reliable by letting the server give permissions to smartphones to produce ultrasound signals. This way positioning requests can be queued and processed one at a time, avoiding overlaps. Communication between one of the phones and the server can go as follows:

- When a user initiates a positioning procedure, his smartphone sends a message to the server over Wi-Fi requesting permission to send ultrasound signal.
- The server receives the message, takes note of the phone's IP address and adds it to the queue.
- When the given entry is reached in the queue, the server sends a permission.
- The phone receives the permission and immediately sends an ultrasound signal.
- The server analyzes the data collected from the microphones, generates the most likely coordinates of the phone and sends them back to the phone over Wi-Fi.
- The phone receives the coordinates, updates the screen and notifies the user with vibration or a sound.

The biggest disadvantage of this method is the arbitrary delay between positioning initialisation and completion. However the extent of this problem largely depends on how the system will be used. For example, if there are usually only a few users in the room that initiate positioning relatively infrequently, it is unlikely that the queue has any entries at all, and the request will be processed with little to no delay. In this case the queue should be seen as a mere precaution. A more robust approach suitable for a higher concentration of users and requests is presented below.

**Multi-frequency signals.** In order for the positioning system to listen for and detect signals from several devices simultaneously, it must be able to distinguish signals coming from different devices. The two biggest limitations are poor detection of changes in volume and a very narrow frequency range. The first limitation mainly means that delivering a unique identifier using only changes in volume is not going to work well. The second constraint means frequencies used in the signal have to fit in the 20-22 kHz range. Practically only about 4 frequencies can be used in this band together, without each setting off a neighbouring filter. While it is possible to tell apart signals packed much closer together using a spectrogram, this is not something that will work well with real-time signal processing. Fortunately there is one more usable parameter – delay. We propose using a combination of 4 different frequencies and pauses of various lengths between the signals to uniquely identify a mobile phone in the room. Care has to be taken not to use the same frequency in succession to avoid problems arising from multipath.

### 6.4.3. Custom beacons

Implementation of the positioning system as a set of microphones connected via a 4-channel audio card to a PC running positioning software should be regarded as a proof of concept. The cost of equipment used as infrastructure (4 microphones, 1 sound card, 1 laptop) for experiments described in this thesis is around 8000 euro. This number could be easily halved by using cheaper microphones. However infrastructure for the positioning approach described in this thesis ideally should be implemented in the form of custom built hardware. Four microphone modules and one computational module should be sufficient to enable positioning in areas close to 7 by 7 meter dimensions or smaller. Computational module can be either placed together with one of the microphones or in a separate casing. Each separate microphone can be placed in a corner and connected to the computational module with cables. Connecting microphones wirelessly doesn't give any advantages, as microphones will have to be connected to a power supply in that case. Many modern public buildings have dropped ceilings, which makes installing microphone modules and hiding wiring and other components very easy. Also in a perfectly rectangular room calibrating the system can be as simple as providing dimensions of the room, provided that every microphone can be placed precisely in the corner.

As an example the following hardware can be used. Knowles SPM0204UD5<sup>32</sup> ultrasonic acoustic sensor can be used as a microphone. Arndale Board<sup>33</sup> can be used as

---

<sup>32</sup> "Ultrasonic Acoustic Sensor" Retrieved 10 January, 2013, from  
<http://www.farnell.com/datasheets/318029.pdf>

the basis of the computational module. Unfortunately it doesn't accept 4-channel audio input, however thanks to its modular structure it should be possible to replace the default audio module with one that has four channels. The board has a Wi-Fi module and a GPU, which can be effectively utilised for matrix manipulations. Even if audio upgrade raises the price of the board by 20%, together with casing, wiring, Wi-Fi antenna and power supply, altogether the setup shouldn't cost more than 300 euro. Larger rooms can be covered by several separate systems placed side by side. Ability to seamlessly transfer connection with the phone from one access point to another will be necessary, which is easier done with Bluetooth.

#### 6.4.4. Signal Reception Model

A polar contour plot for ultrasound energy propagation was introduced on Figure 26 in Section 3.3. The plot roughly resembles a cardioid and can be used to estimate how well signal reception will be at a certain distance from the speaker, and angle from the direction the speaker is pointing, provided that the microphone points directly at the speaker. Considering the orientation of microphones is fixed, most of the time they will not be facing the mobile phone directly and the angle will vary depending on the user's location. How well a microphone can detect sound from a particular angle and distance is also traditionally depicted using polar plots which often resemble cardioids. Many professional microphones come with a polar plot supplied by the manufacturer. Unfortunately they only provide a plot for one intensity level which is enough to only predict an overall shape of the propagation model. A plot needs to have multiple layers

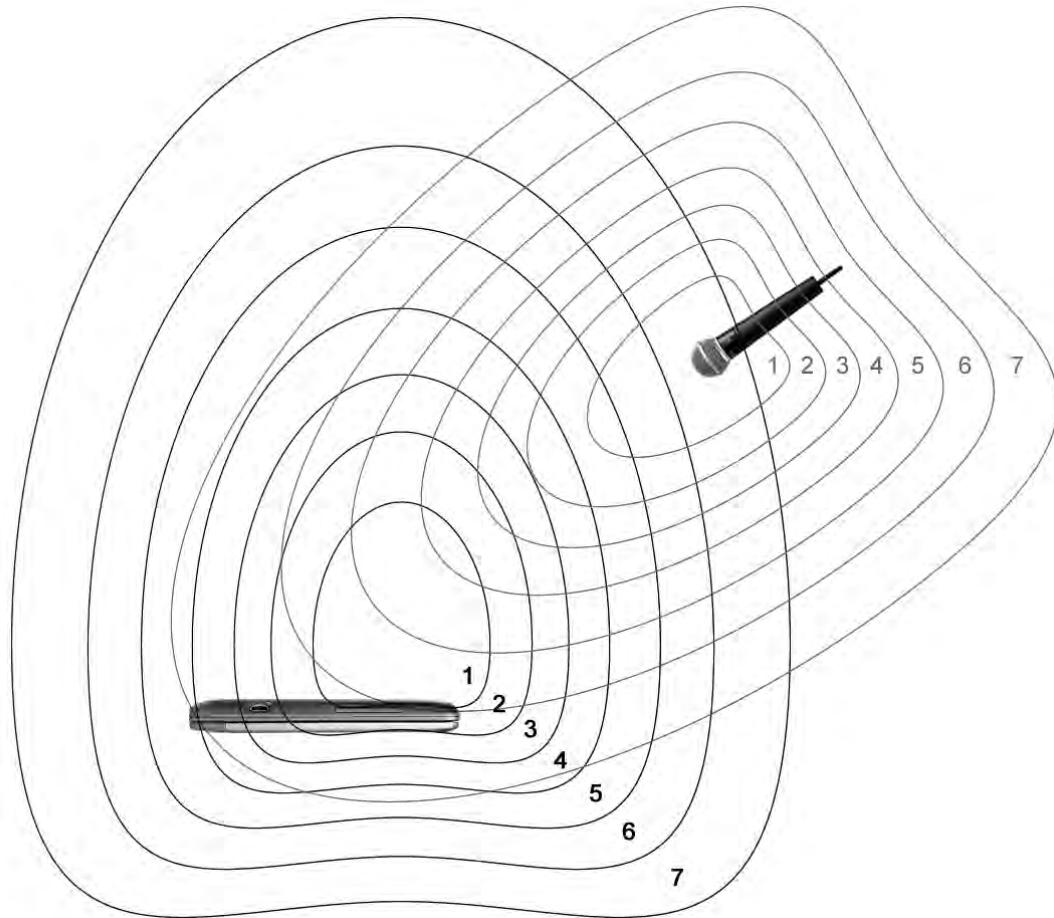
---

<sup>33</sup> "Arndale Board" Accessed 10 January, 2013, at  
[http://www.arndaleboard.org/wiki/index.php/Main\\_Page](http://www.arndaleboard.org/wiki/index.php/Main_Page)

of reception intensity so that it can be used to predict reception quality at a particular angle and distance.

We suspect that it can be effectively predicted how well an ultrasound signal will be detected based on the combination of two factors. Which layer of the phone's polar plot the microphone intersects, and which layer of the microphone's plot the phone intersects. See Figure 51. Overall reception quality for a microphone and a phone placed in a particular way in relation to each other will most likely be an average of these two variables, provided that each layer of the plot was assigned a number corresponding to its intensity. It is possible that some mathematical relationship, other than average, more accurately represents reception quality. For a positioning scenario with four microphones, positioning quality at a particular orientation and position in the room will be the worst of the four averages corresponding to reception quality between the phone and each individual microphone.

If this model happens to be true, it will be possible to predict the behaviour of an ultrasound positioning system with some degree of precision. For example identify dead zones and make amendments in the microphone layout, give the user an accurate confidence factor for each individual positioning fix, introduce weights to the trilateration procedure and overall make the system more predictable and robust.



**Figure 51: Model of reception quality.** In this example a phone's speaker intersects layer 6 of the microphone's polar plot. Microphone intersects layer 7 of the speaker's polar plot.

In order to test this hypothesis, we propose the following procedure:

1. Take measurements and generate a polar plot for one of the microphones in a similar fashion to how a polar plot was made for the mobile phone in Section 3.3.
2. Make a 3D model of the room, place microphones in the correct positions and wrap the polar plots 360 degrees around them in such a way that they correctly represent reception quality.
3. Place the phone together with its polar plot wrapped around the speaker into different locations in the 3D model. Calculate which layer of the microphone's plot

the phone intersects with and which layer of the phone's plot the microphone intersects with individually for each of the four microphones. Estimate positioning quality using this data and compare to either new or existing experimental data. Adjust the estimation process until estimated and experimental results match consistently.

4. Use the computer estimation model for a completely new location and microphone layout to verify that the model is versatile.

## 6.5. Overall Conclusions

This thesis describes design, testing and evaluation of an ultrasound indoor positioning method for off-the-shelf smartphones. Ultrasound trilateration was identified as a very promising approach. Signals travelling at the speed of sound offer very good accuracy, well below one metre even with standard sound hardware. Also the 20-22 kHz range is normally inaudible to humans but can be reproduced by mobile phone speakers and captured with standard microphones. Also because no specialised hardware is used on the client side, if the user has a smartphone s/he only needs to make a software installation to be able to use the positioning system. The highly directional nature of ultrasound, susceptibility to certain noises, and the need for line-of-sight between speaker and receiver were identified as the biggest obstacles to positioning accuracy.

A prototype of the positioning system was developed and tested for accuracy and possible shortcomings. Despite obstacles listed above it was possible to get full coverage of a 7 by 7 metre room with four microphones placed directly below the ceiling in the corners of the room and achieve a certain degree of reliability. On average the system had accuracy of around 10 centimetres, an order of magnitude better than

contemporary approaches. In the given implementation it is desirable for the user to flip the phone upside down (speaker up) to get the best accuracy, but otherwise the system behaves exactly like an on-demand positioning system is expected to behave. There are occasional problems with accuracy when a line-of-sight is blocked or the user is outside the optimal reception space of one of the microphones. This problem can be addressed by adding more microphones as well as using omnidirectional microphones.

Ultrasound indoor positioning is a very promising approach particularly because there are no other technologies available for off-the-shelf mobile phones that can offer real-time indoor positioning with comparable accuracy. It can therefore be seen as a potential positioning platform for indoor location based services, currently an emerging market, after some more research into reliability, scalability, and mass production and deployment is carried out.

## 7. REFERENCES

- Addlesee, M., Curwen, R., Hodges, S., Newman, J., Steggles, P., Ward, A., Hopper, A. (2001). "Implementing a Sentient Computing System." IEEE Computer **34**(8): 50-56.
- Anderson, R., Bilger, H. R., Stedman, G. E. (1994). "„Sagnac“ effect: A century of Earth-rotated interferometers." American Journal of Physics **62**(11): 975-985.
- Arrington, M. (2009, 9 November). "Google Redefines GPS Navigation Landscape: Google Maps Navigation For Android 2.0." TechCrunch, from <http://www.techcrunch.com/2009/10/28/google-redefines-car-gps-navigation-google-maps-navigation-android/>.
- Aubeck, F., Isert, C., Gusenbauer, D. (2011). Camera based step detection on mobile phones. IPIN. Guimarães, Portugal, IEEE: 1-7.
- Badea, V., Eriksson, R. (2005). Indoor Navigation with Pseudolites (fake GPS Sat.). Department of Science and Technology. Linköping, Linköping University. **Master of Science**.
- Bahl, P., Padmanabhan, V. (2000). RADAR: An In-Building RF-Based User Location and Tracking System. INFOCOM, IEEE. **2**: 775-784.
- Ball, M. (2012). "Sensors & Systems, Google Has a Strong Start on the Indoor Location Frontier." Retrieved 6 November, 2012, from <http://www.sensysmag.com/dialog/interviews/28563-google-has-a-strong-start-on-the-indoor-location-frontier.html>.
- Bancroft, S. (1985). An algebraic solution of the GPS equations. IEEE Transactions on Aerospace and Electronic Systems. **21**: 56-59.
- Banks, K. (2002). "The Goertzel Algorithm." Retrieved 8 January, 2013, from <http://www.embedded.com/design/configurable-systems/4024443/The-Goertzel-Algorithm>.
- Baunach, M., Kolla, R., Mühlberger, C. (2007). SNOW Bat: A high precise WSN based location system, Universität Würzburg, Lehrstuhl für Informatik V.
- Borio, D., O'Driscoll, C., Fortuny-Guasch, J. (2011). Pulsed Pseudolite Signal Effects on Non-Participating GNSS Receivers. IPIN. Guimarães, Portugal, IEEE: 1-6.
- Borriello, G., Liu, Alan., Offer, T., Palistrant, C., Sharp, R. (2005). WALRUS: Wireless Acoustic Location with Room-Level Resolution using Ultrasound. Mobisys: 191-203.
- Bossler, J., Jensen, J., McMaster, R., Rizos, C. (2002). Manual of Geospatial Science and Technology, Taylor & Francis.

- Bowditch, N. (1995). Dead Reckoning. The American Practical Navigator: An epitome of navigation.
- Bres, S., Tellez, B. (2009). Localisation and Augmented Reality for Mobile Applications in Cultural Heritage 3rd ISPRS International Workshop.
- Cheung, K., Intille, S., Larson, K. (2006). An inexpensive Bluetooth-based indoor positioning hack. UbiComp.
- Chou, L., Lee, C., Lee, M., Chang, C. (2004). A Tour Guide System for Mobile Learning in Museums. 2nd IEEE International Workshop on Wireless and Mobile Technologies in Education, IEEE: 195-196.
- Cobb, S. (1997). GPS Pseudolites: Theory, Design, and Application. Department of Aeronautics and Astronautics. Palo Alto, Stanford University. **Doctor of Philosophy**.
- de Vries, G., van Beuningen, G. (1997). "Concepts and applications of directivity controlled loudspeaker arrays." The Journal of the Acoustical Society of America **101**(5).
- Dixon, W. (1983). BMDP statistical software, University of California Press.
- Egenhofer, M. (1999). Spatial Information Appliances: A next Generation of Geographic Information Systems. Geo-info.
- Ferris, B., Hähnel, D., Fox, D (2006). Gaussian Processes for Signal Strength-Based Location Estimation. Robotics Science and Systems
- Foley, J., Dam, A., Feiner, S., Hughes, J. (1996). Computer Graphics: Principles and Practice in C, Addison-Wesley Professional.
- Ghilani, C., Wolf, P. (2006). Adjustment Computations: Spatial Data Analysis, John Wiley & Sons, Inc.
- Gold, R. (1967). Optimal binary sequences for spread spectrum multiplexing IEEE Trans. Information Theory. **13**: 619-621.
- Goyal, P., Ribeiro, V., Saranz, H., Kumar, A. (2011). Strap-Down Pedestrian Dead-Reckoning System. IPIN. Guimarães, Portugal, IEEE: 1-7.
- Hallberg, J., Nilsson, M., Synnes, K. (2003). Positioning with Bluetooth. 10th International Conference on Telecommunications, IEEE. **2**: 954-958.
- Harter, A., Hopper, A., Steggles, P., Ward, A., Webster, P. (1999). The Anatomy of a Context-Aware Application. Mobile Computing and Networking: 59-68.
- Hazas, M., Hopper, A. (2006). Broadband Ultrasonic Location Systems for Improved Indoor Positioning. IEEE Transactions on Mobile Computing, IEEE. **5**: 536-547.
- Hoene, C., Willmann, J. (2008). Four-way TOA and software-based trilateration of IEEE 802.11 devices Personal, Indoor and Mobile Radio Communications, 2008. Cannes, France, IEEE: 1-6.

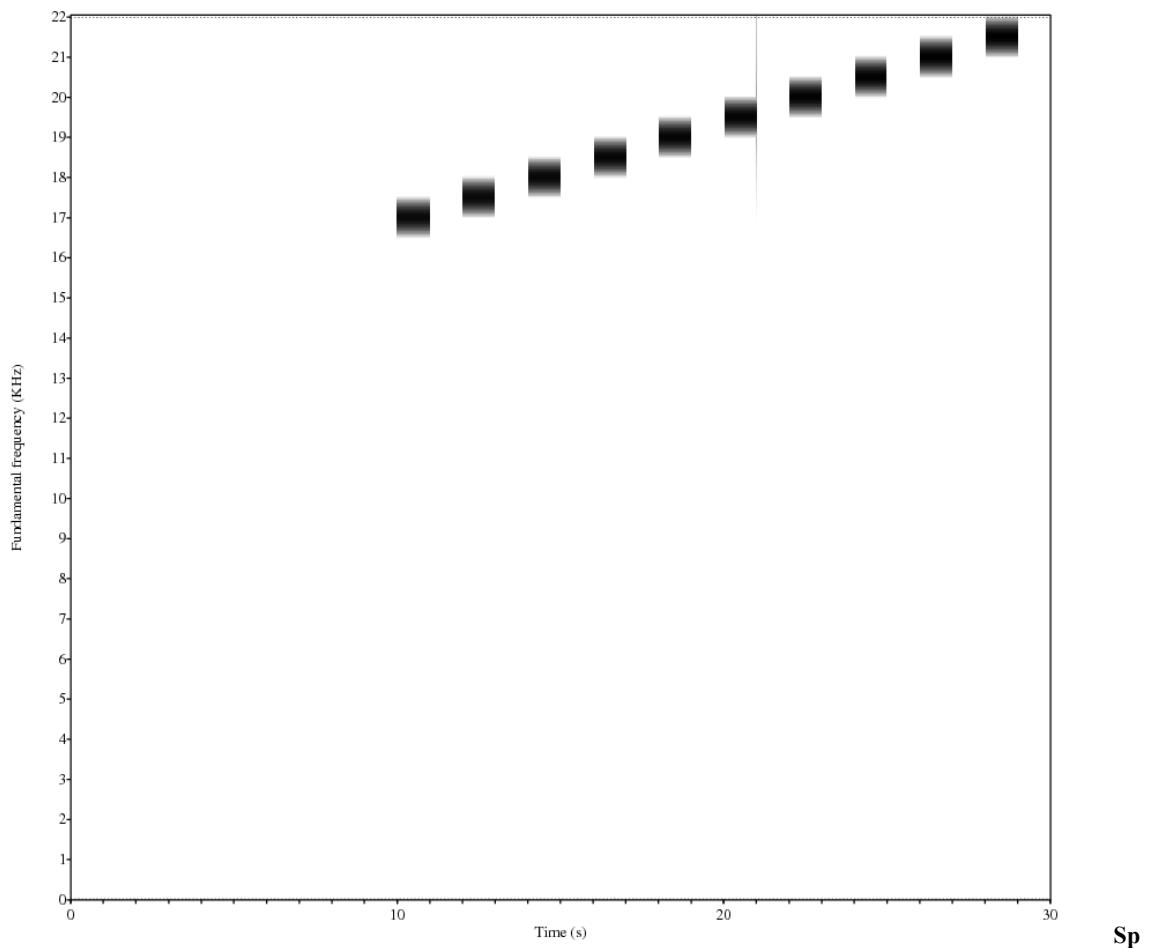
- Holm, S. (2009). Hybrid Ultrasound–RFID Indoor Positioning: Combining the Best of Both Worlds. 2009 IEEE International Conference on RFID, IEEE: 155-162.
- Kolodziej, K., Hjelm, J. (2006). Local positioning systems : LBS applications and services.
- Kratz, S., Ballagas, R. (2007). Gesture recognition using motion estimation on mobile phones. 3rd International Workshop on Pervasive Mobile Interaction Devices (PERMID'07). Toronto, Ontario, Canada.
- Krumm, J., Horvitz, E. (2004). LOCADIO: inferring motion and location from Wi-Fi signal strengths Mobile and Ubiquitous Systems: Networking and Services, IEEE: 4-13.
- Liu, Y., Wilde, E. (2011). Personalized location-based services. Proceedings of the 2011 iConference: 496-502
- Lowe, D. G. (1999). Object recognition from local scale-invariant features Computer Vision, IEEE. **2**: 1150-1157.
- Maddio, S., Bencini, L., Cidronali, A., Manes, G. (2010). A Single Anchor Direction of Arrival Positioning System Augmenting Standard Wireless Communication Technology. IPIN 2010. Zurich, Switzerland: 19-20.
- Madhavapeddy, A., Scott, D., Sharp, R. (2003). Context-Aware Computing with Sound. 5th International Conference on Ubiquitous Computing: 315-332.
- Meijers, M., Zlatanova, S., Pfeifer, N. (2005). 3D Geo-Information Indoors: Structuring for Evacuation. First International Workshop on Next Generation 3D City Models: 11-16.
- Mestre, P., Serodio, C., Coutinho, L., Reigoto, L., Matias, J. (2011). Hybrid technique for Fingerprinting using IEEE802.11 Wireless Networks, Guimarães, Portugal, IEEE.
- Meyer, D. (2009). "Bluetooth 3.0 released without ultrawideband." Retrieved 8 January, 2013, from <http://www.zdnet.com/bluetooth-3-0-released-without-ultrawideband-3039643174/>.
- Minami, M., Fukuju, Y., Hirasawa, K., Yokoyama, S., Mizumachi, M., Morikawa, H., Aoyama, T. (2004). "DOLPHIN: A Practical Approach for Implementing a Fully Distributed Indoor Ultrasonic Positioning System." Lecture Notes in Computer Science Volume 3205: 347-365.
- Modsching, M., Kramer, R., Hagen, K. t. (2006). Field trial on GPS Accuracy in a medium size city: The influence of built-up. WPNC. Hannover, Germany: 209-218.
- Nakazato, Y., Kanbara, M., Yokoya, N. (2005). Localization of Wearable Users Using Invisible Retro-reflective Markers and an IR Camera. SPIE proceedings. **5664**: 563-570.

- O'Connor, M. (1997). Carrier-Phase Differential GPS for Automatic Control of Land Vehicles. Department of Aeronautics and Astronautics, Stanford University. **Doctor of Philosophy**.
- Otsason, V., Varshavsky, A., LaMarca, A., De Lara, E. (2007). Accurate GSM Indoor Localization. Pervasive and Mobile Computing. 3.
- Paarmann, L. (2001). Design and Analysis of Analog Filters: A Signal Processing Perspective, Springer.
- Packi, F., Beutler, F., Hanebeck, U. (2010). Wireless Acoustic Tracking for Extended Range Telepresence. IPIN. Zurich, Switzerland, IEEE: 1-9.
- Pals, H., Dai, Z., Grabowski, J., Neukirchen, H. (2003). UML-Based Modeling of Roaming with Bluetooth Devices. Proceedings of the First Hangzhou-Lübeck Workshop on Software Engineering University of Hangzhou, China, University of Hangzhou, China.
- Peng, C., Shen, G., Zhang, Y., Li, Y., Tan, K. (2007). BeepBeep: A High Accuracy Acoustic Ranging System using COTS Mobile Devices. SenSys: 1-14.
- Perez, M. (2012, November 29th, 2011). "Nokia shows off super accurate, in-door 3D mapping." Retrieved 17 January, 2012, from [www.intomobile.com/2011/11/29/nokia-shows-off-super-accurate-3d-indoor-mapping/](http://www.intomobile.com/2011/11/29/nokia-shows-off-super-accurate-3d-indoor-mapping/).
- Priyantha, N. (2005). The Cricket Indoor Location System. Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology. **Doctor of Philosophy in Computer Science and Engineering**.
- Randell, C., Djiallis, C., Muller, H. (2003). Personal Position Measurement Using Dead Reckoning. 7th IEEE International Symposium on Wearable Computers, IEEE: 166-173.
- Randell, C., Muller, H. (2001). Low Cost Indoor Positioning System. Ubicomp 2001: Ubiquitous Computing. G. D. Abowd: 42-48.
- Rivington, J. (2012). "techradar.av, Project Glass: what you need to know." Retrieved 8 August, 2012, from <http://www.techradar.com/news/video/project-glass-what-you-need-to-know-1078114>.
- Rumsey, F. (2009). Sound and Recording, Focal Press.
- Ruotsalainen, L., Kuusniemi, H., Chen, R. (2011). Heading Change Detection for Indoor Navigation with a Smartphone Camera. IPIN. Guimarães, Portugal, IEEE: 1-7.
- Scarfone, K., and Padgett, J. (2008). Guide to Bluetooth Security, National Institute of Standards and Technology. <http://csrc.nist.gov/publications/nistpubs/800-121/SP800-121.pdf>.
- Schiller, J., Voisard, A. (2004). Location-Based Services, Morgan Kaufmann.

- Siciliano, B., Khatib, O. (2008). Springer Handbook of Robotics, Springer.
- Simon, R., Fröhlich, P. (2007). A mobile application framework for the geospatial web. 16th international conference on World Wide Web: 381-390.
- Subhan, F., Hasbullah, H. (2009). Designing a Roaming Protocol for Bluetooth Networks. National Postgraduate Conference (NPC) Universiti Teknologi Petronas, Malaysia.
- Thapa, K., Case, S. (2003). An Indoor Positioning Service for Bluetooth Ad Hoc Networks. MICS.
- Tsai, C., Chou, S., Lin, S. (2010). "Location-aware tour guide systems in museums." Scientific Research and Essays 5(8): 714-720.
- Wagner, D., Reitmayr, G., Mulloni, A., Drummond, T., Schmalstieg, D. (2008). Pose Tracking from Natural Features on Mobile Phones. 7th IEEE/ACM International Symposium on Mixed and Augmented Reality, IEEE Computer Society: 125-134.
- Wagner, D., Schmalstieg, D. (2003). First Steps Towards Handheld Augmented Reality. 7th IEEE International Symposium on Wearable Computers, IEEE Computer Society: 127-136.
- Wang, J., Zhai, S., Canny, J. (2006). Camera Phone Based Motion Sensing: Interaction Techniques, Applications and Performance Study. UIST Montreux, Switzerland: 101-110.
- Wasinger, R., Stahl, C., Krüger, A. (2003). M3I in a Pedestrian Navigation & Exploration System. Human-Computer Interaction with Mobile Devices and Services, Springer Berlin / Heidelberg. 2795/2003: 481-485.
- Williams, B., Klein, G., Reid, I. (2007). Real-Time SLAM Relocalisation. Computer Vision: 1-8.
- Zhou, S., Pollard, J. (2006). "Position measurement using Bluetooth." Consumer Electronics, IEEE Transactions on 52(2): 555-558.

## APPENDIX 1. SPECTROGRAMS

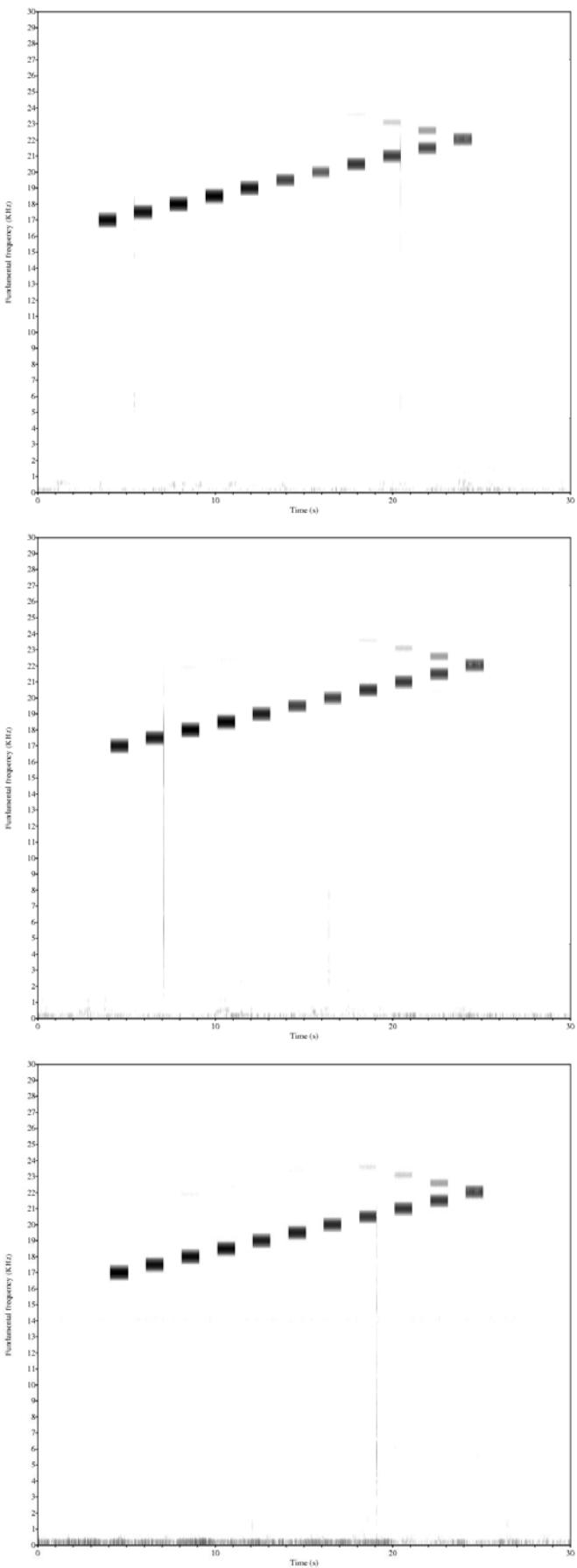
Listed here are spectrograms of recordings made the following mobile phones: HTC G1, HTC Hero, Apple iPhone 3GS and Nokia 6210 Navigator. Spectrogram of the original file that was played back is provided below.



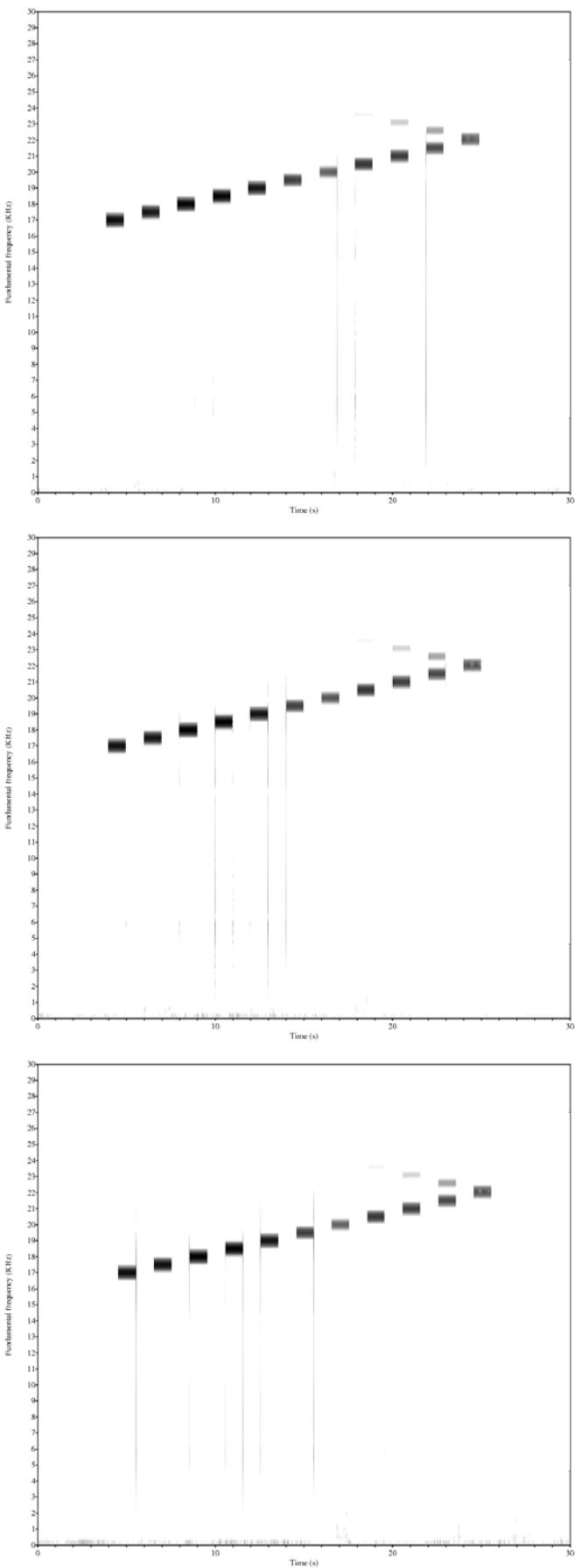
**Spectrogram of the file played back by the smartphones.** X axis depict time and Y axis depict Frequency.

*Chromatic value shows energy.*

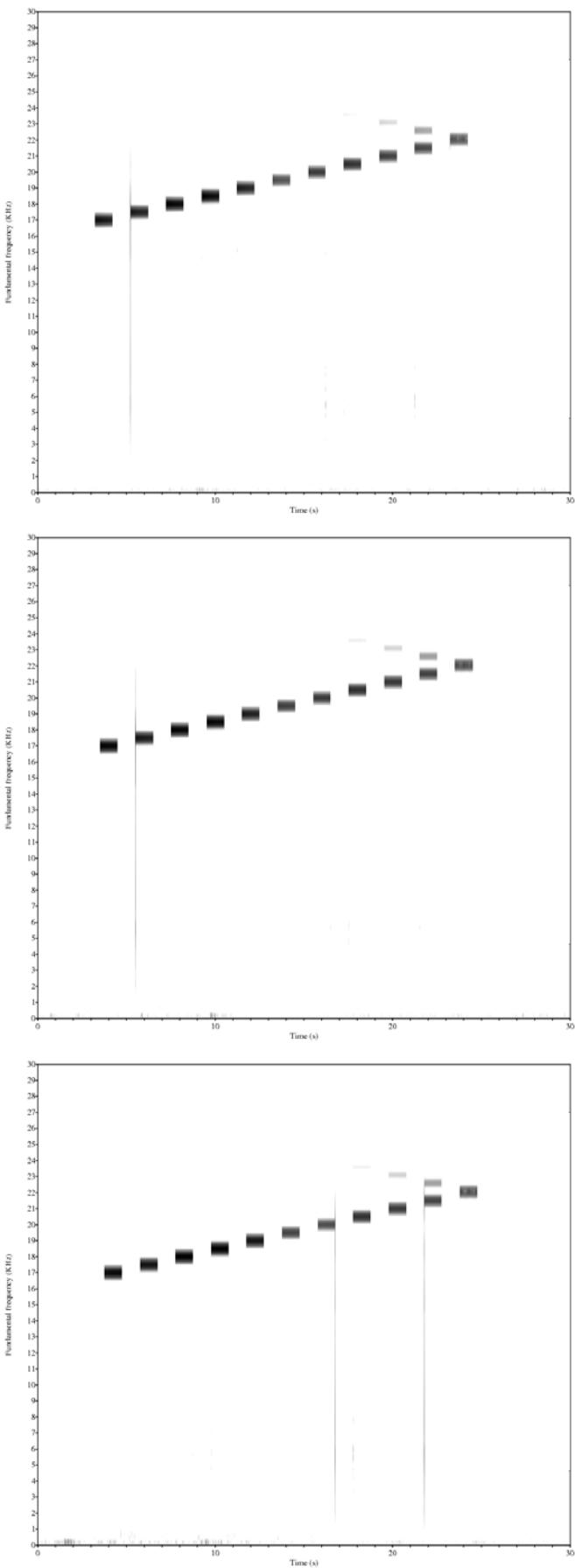
## HTC G1 devphone at 20% file volume



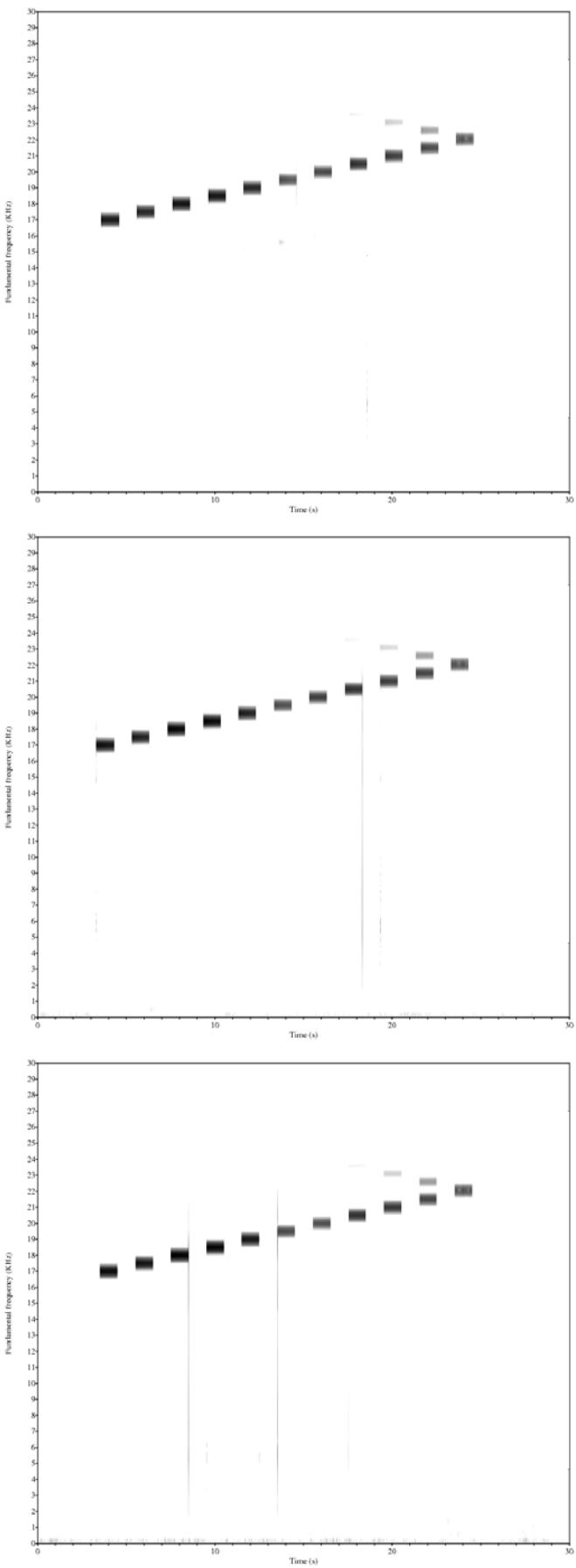
## HTC G1 devphone at 40% file volume



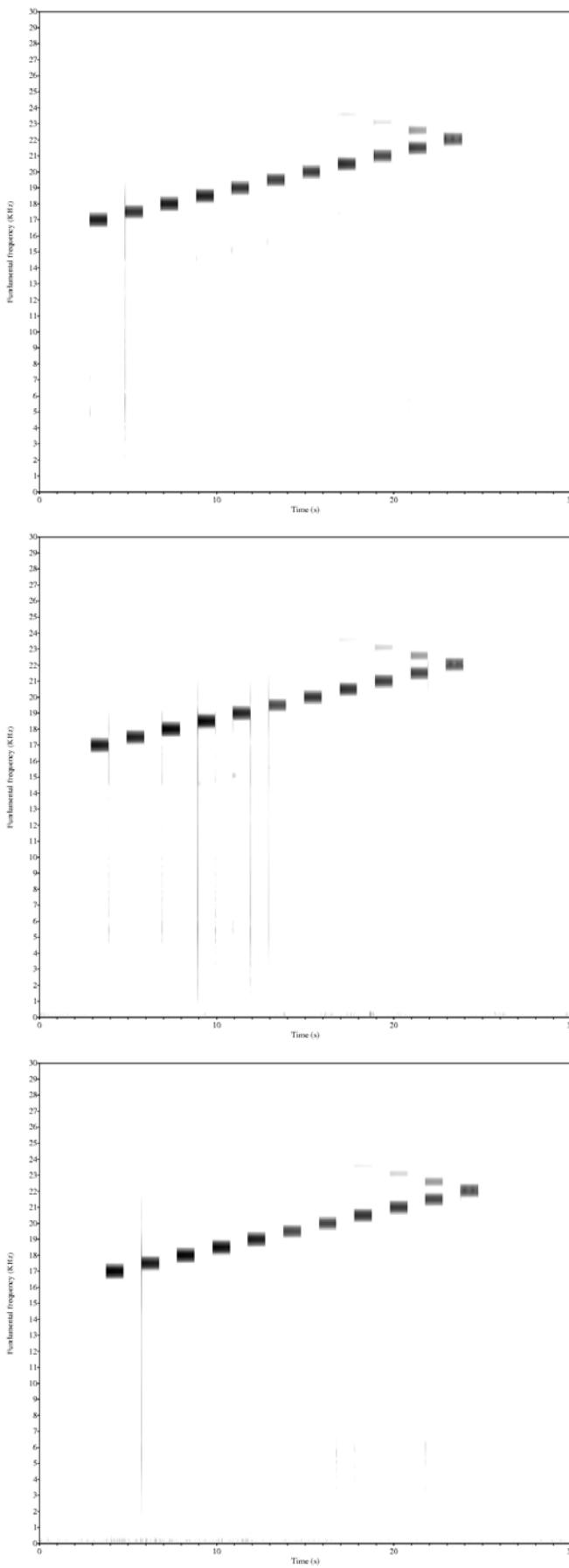
## HTC G1 devphone at 60% file volume

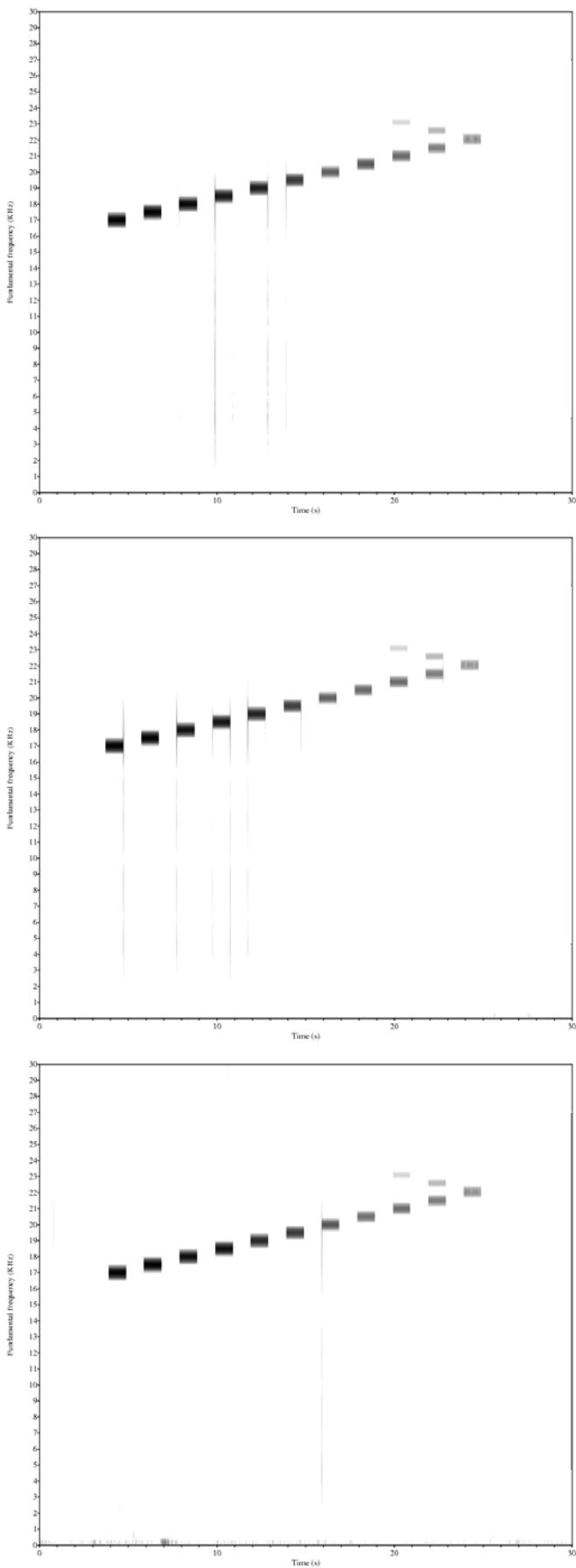


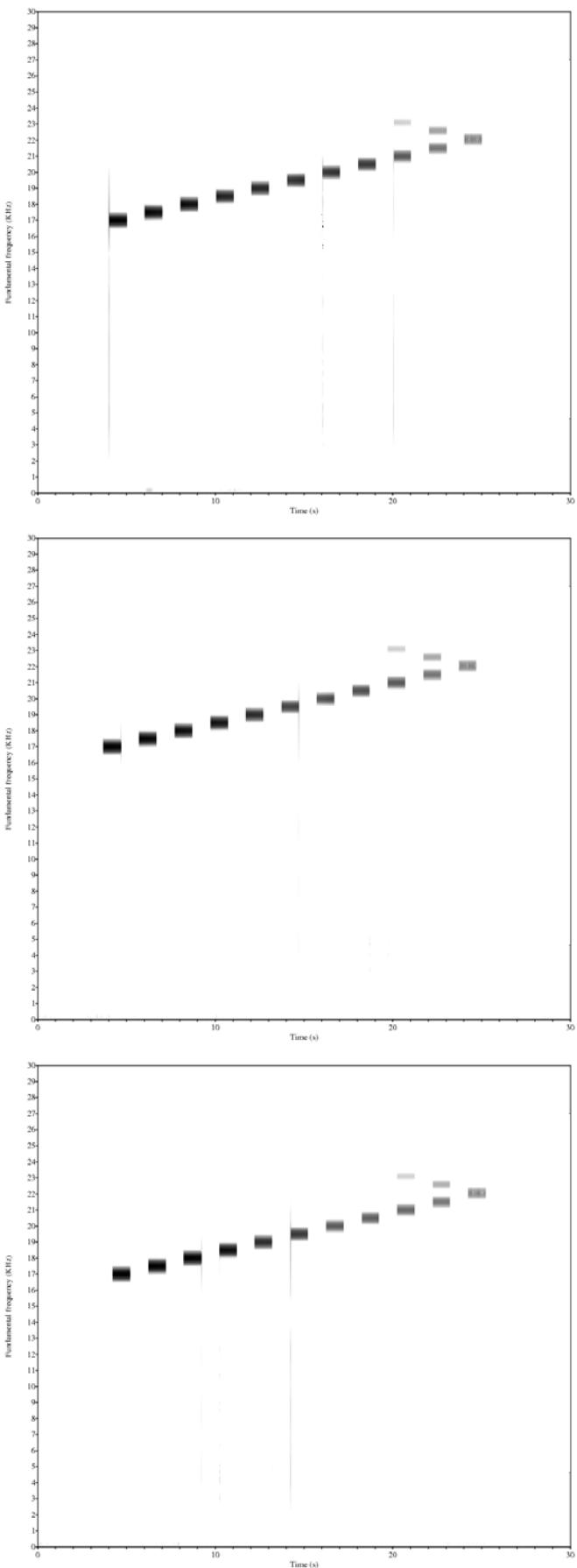
## HTC G1 devphone at 80% file volume

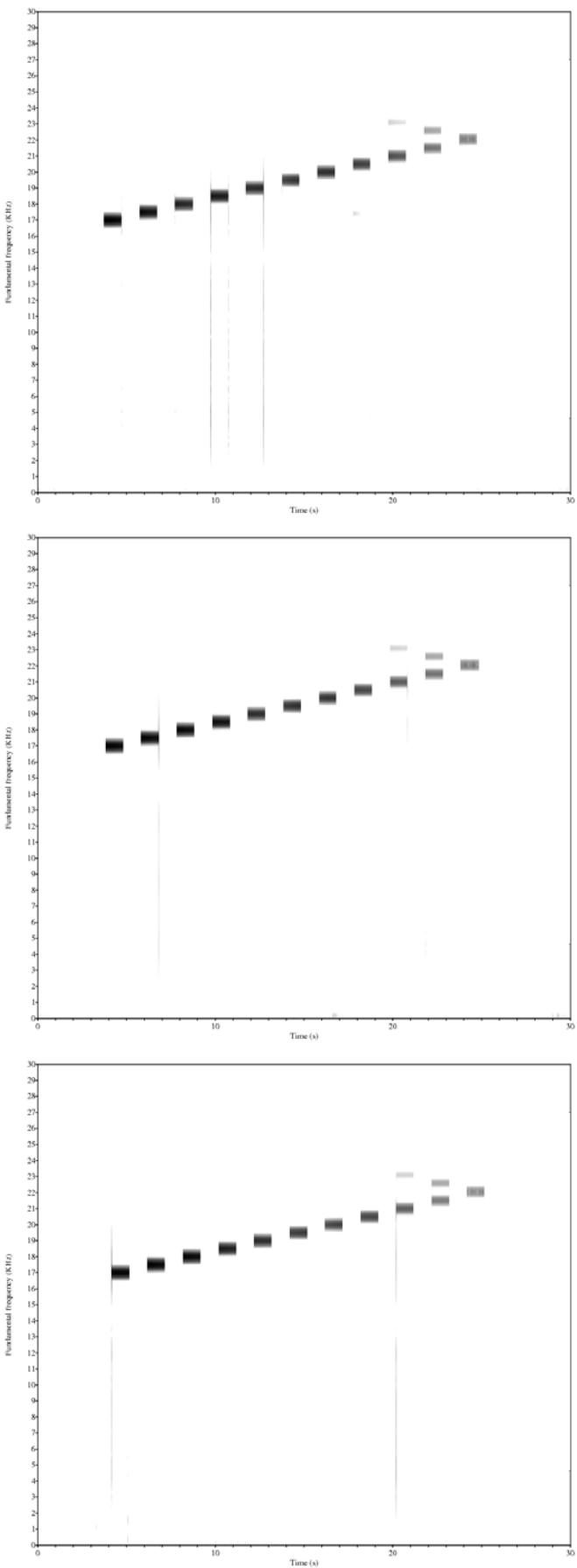


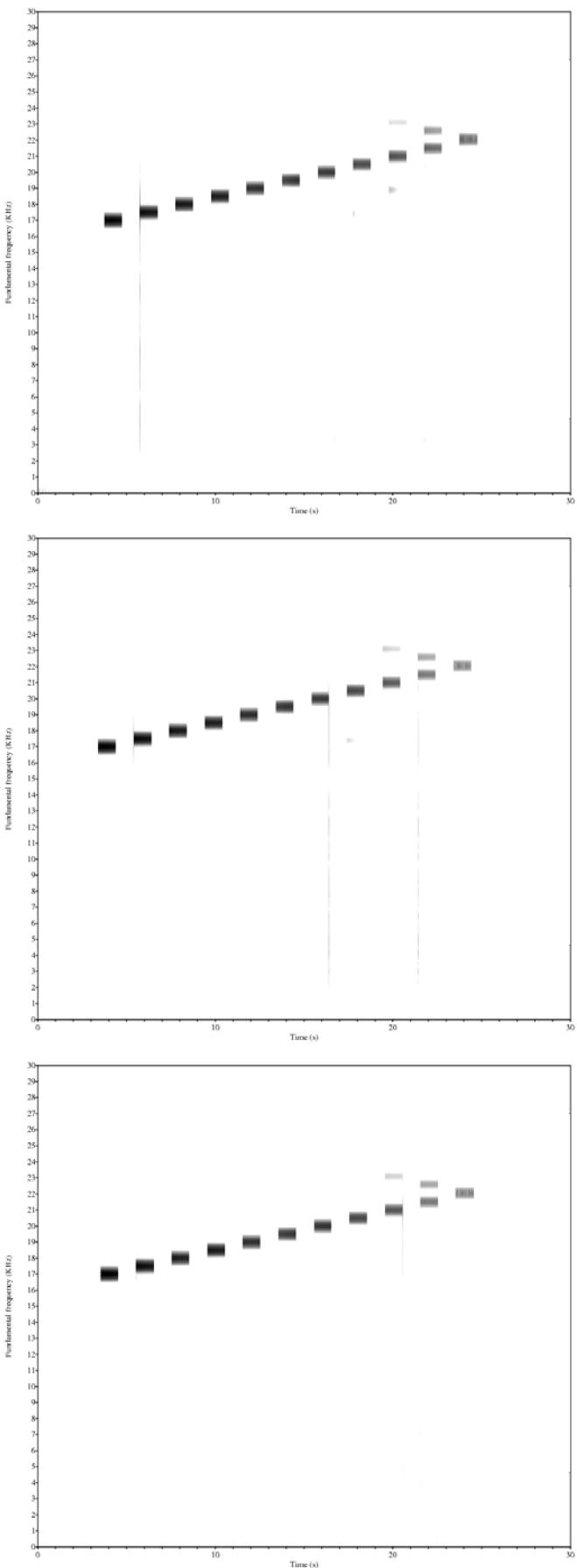
## HTC G1 devphone at 100% file volume

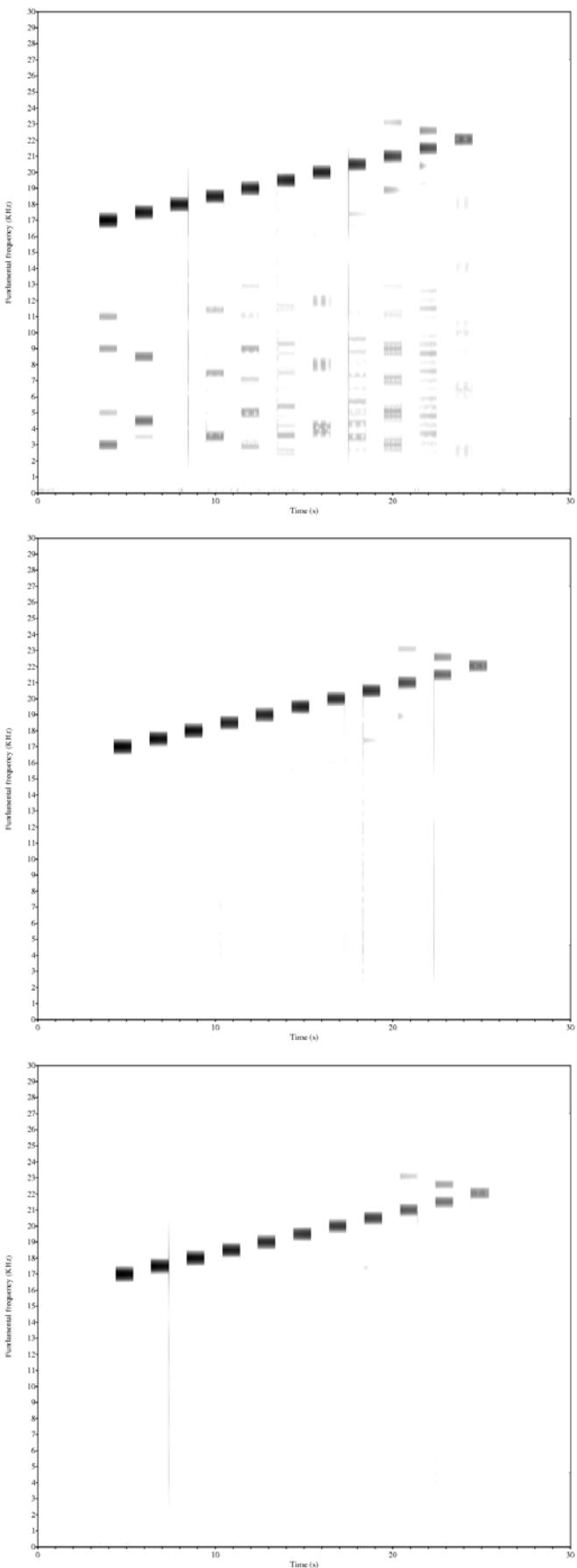


**HTC Hero at 20% file volume**

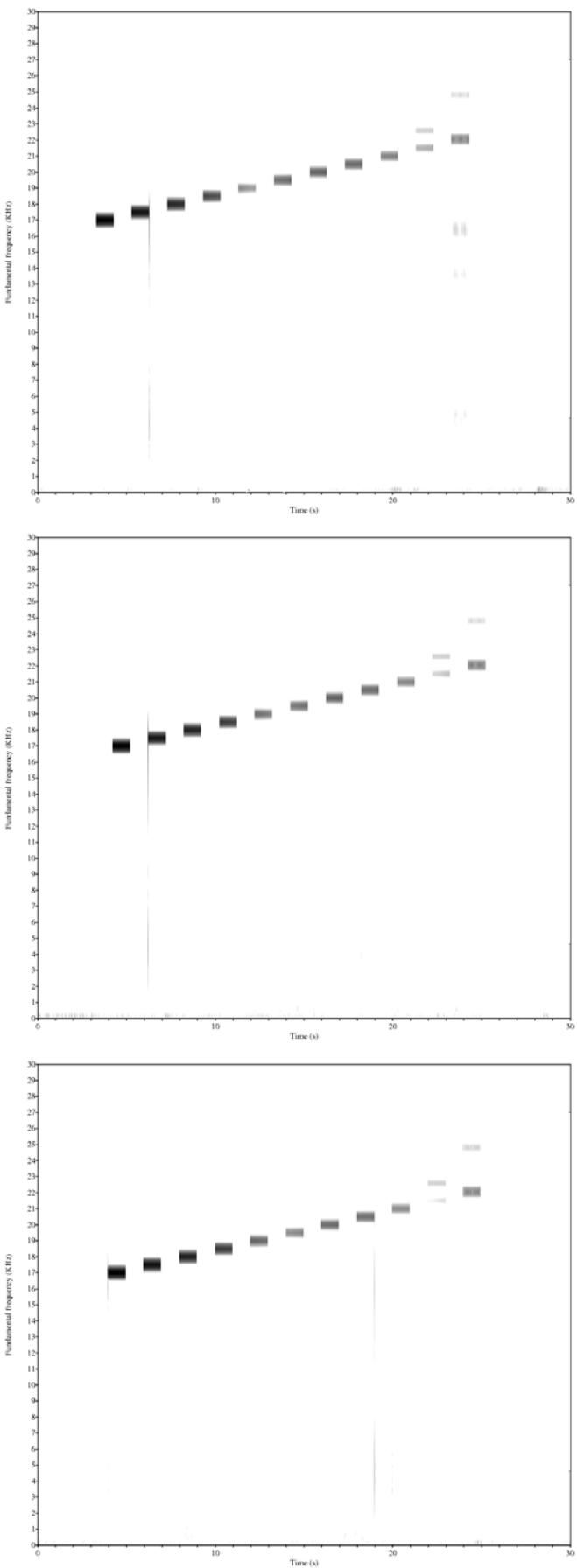
**HTC Hero at 40% file volume**

**HTC Hero at 60% file volume**

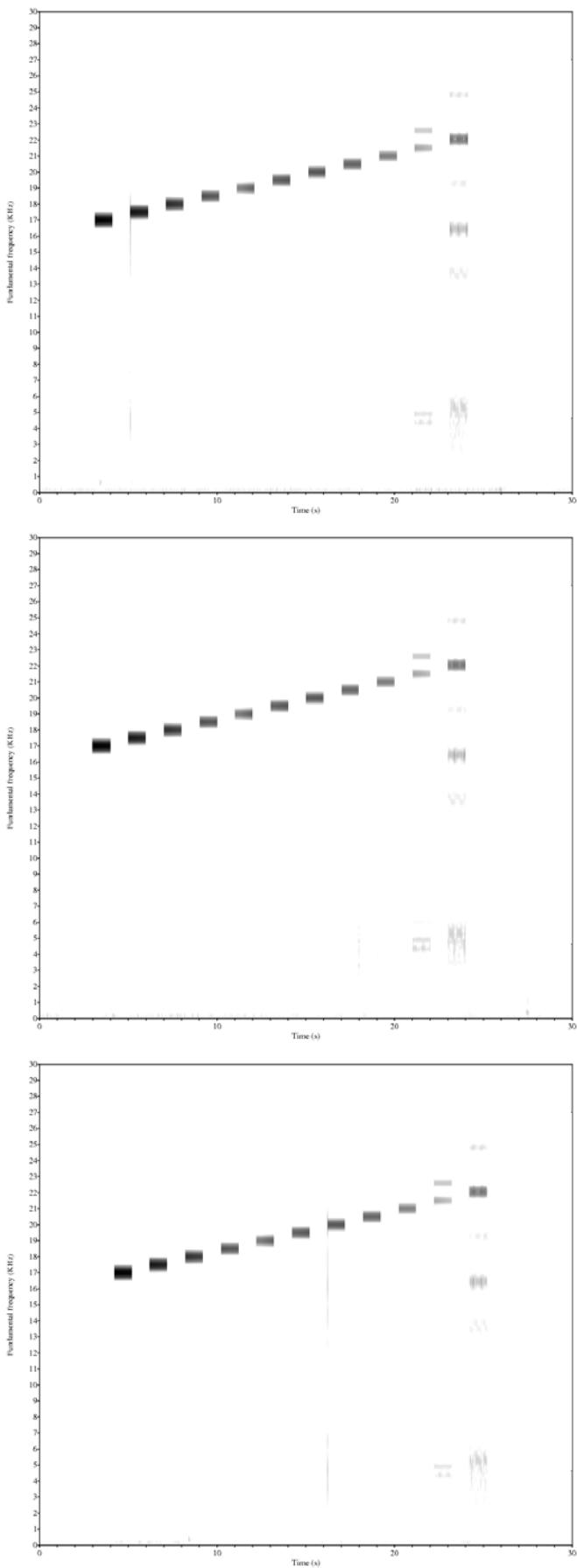
**HTC Hero at 80% file volume**

**HTC Hero at 100% file volume**

## Apple iPhone 3GS at 20% file volume



## Apple iPhone 3GS at 40% file volume

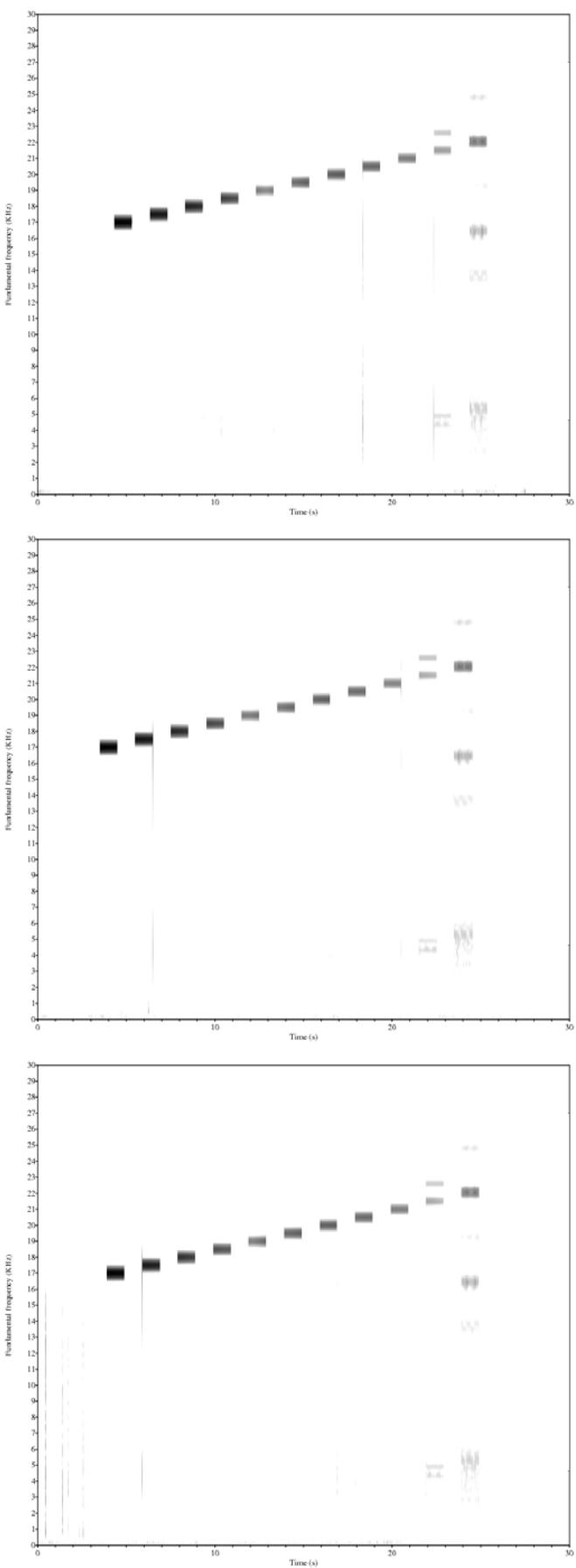


**max volume**

**max volume - 1**

**max volume - 2**

## Apple iPhone 3GS at 60% file volume

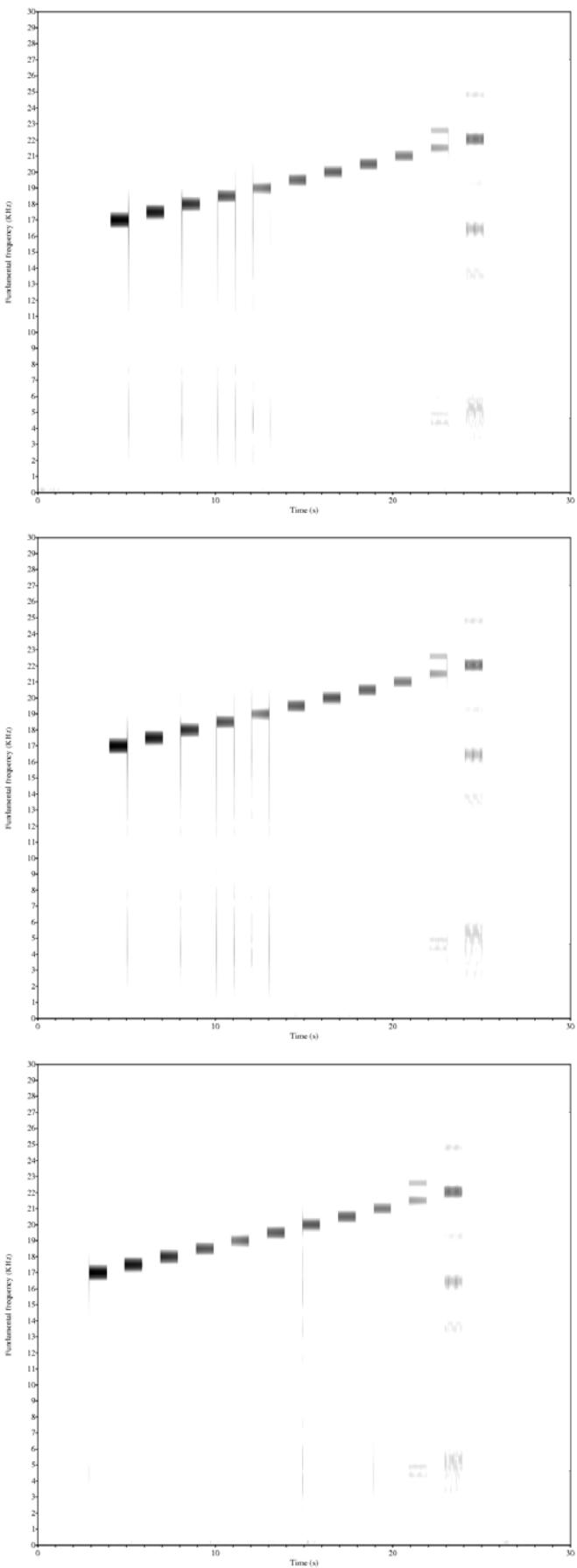


**max volume**

**max volume - 1**

**max volume - 2**

## Apple iPhone 3GS at 80% file volume

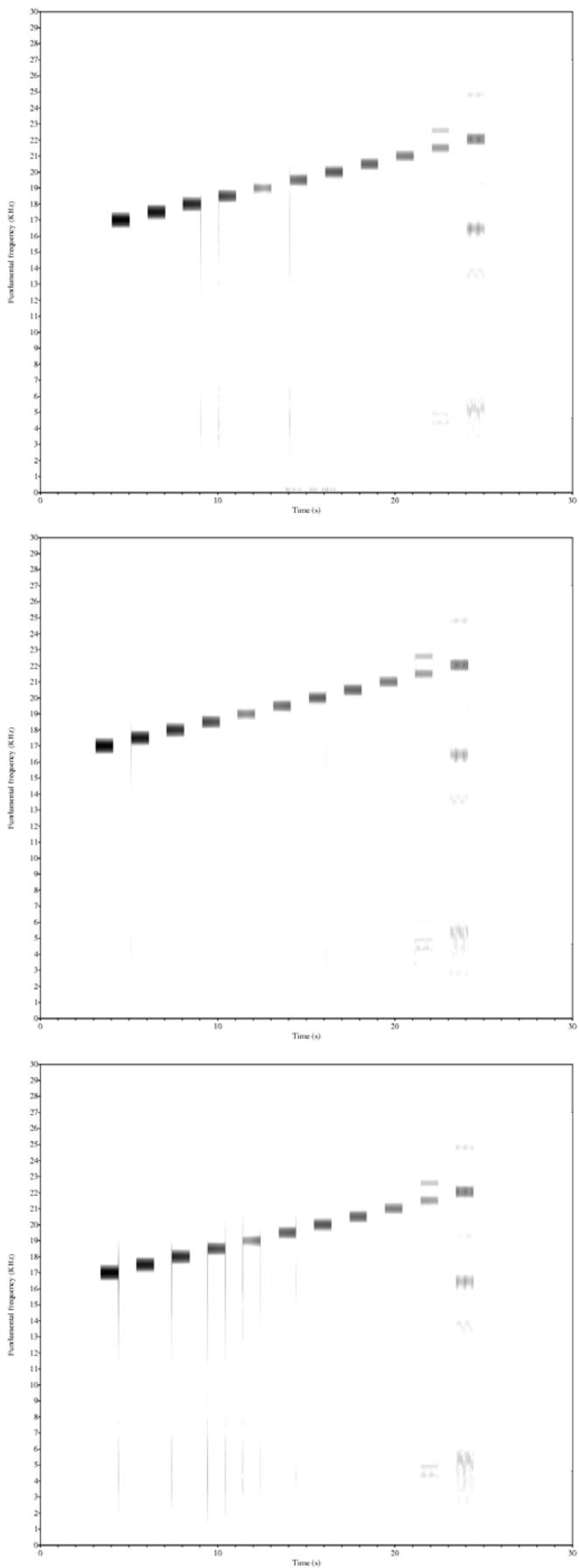


**max volume**

**max volume - 1**

**max volume - 2**

## Apple iPhone 3GS at 100% file volume

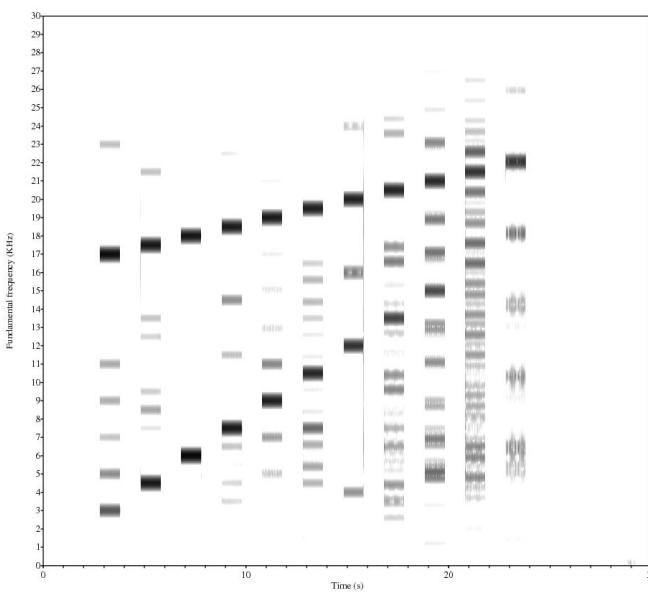


**max volume**

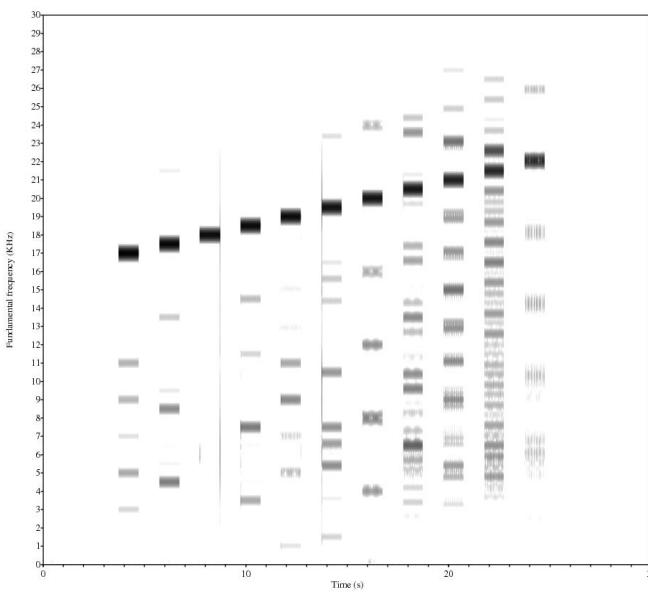
**max volume - 1**

**max volume - 2**

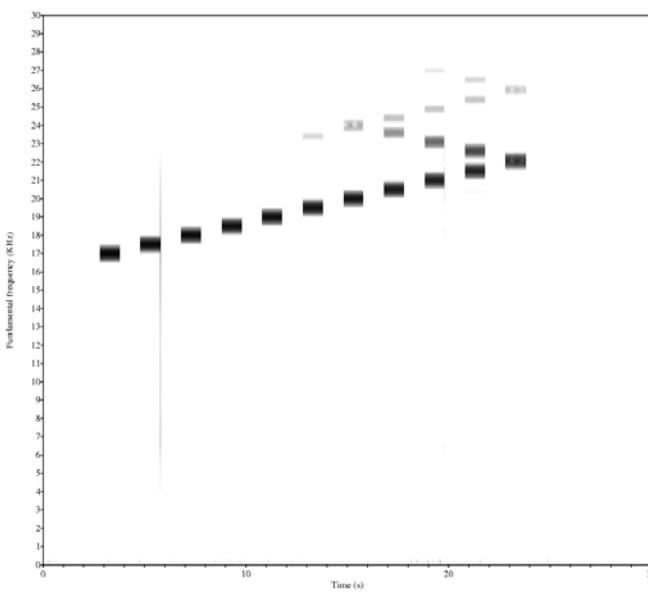
## Nokia 6210 Navigator at 20% file volume



**max volume**

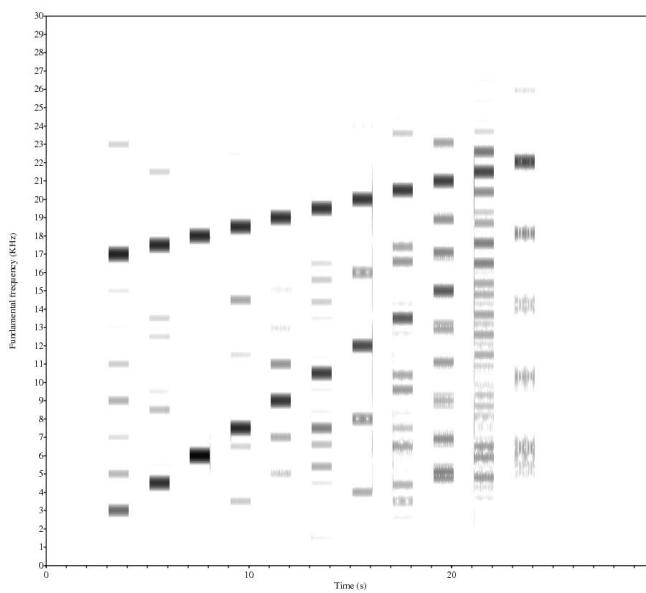


**max volume - 1**

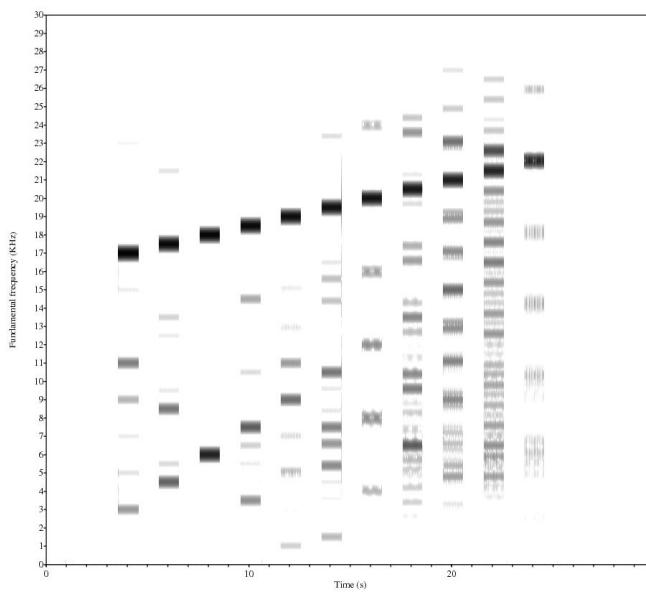


**max volume - 2**

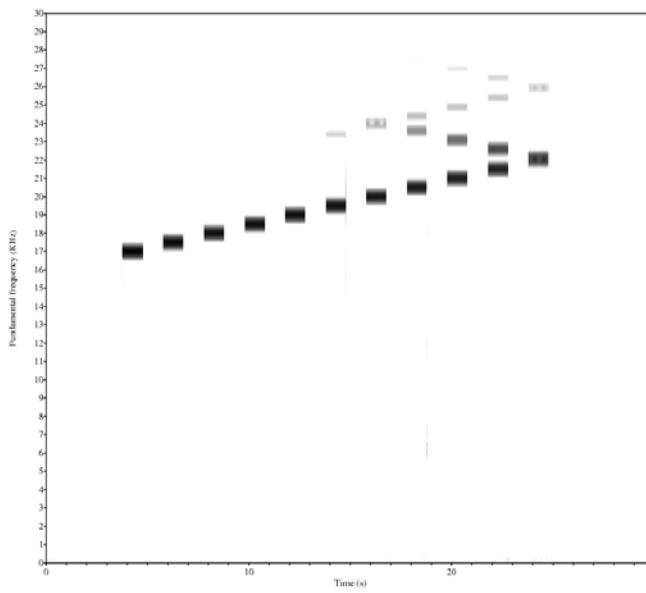
## Nokia 6210 Navigator at 40% file volume



**max volume**

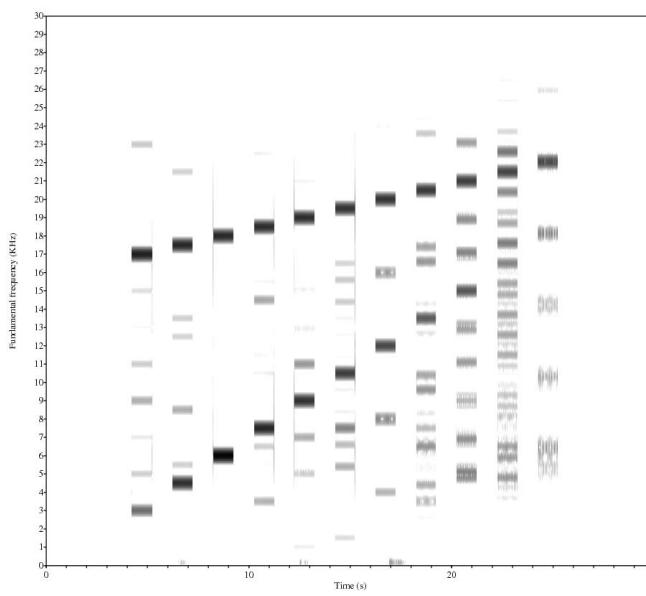


**max volume - 1**

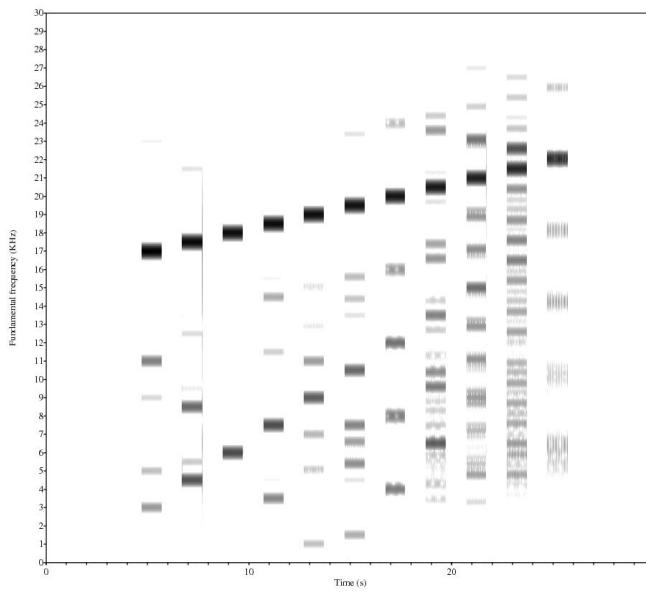


**max volume - 2**

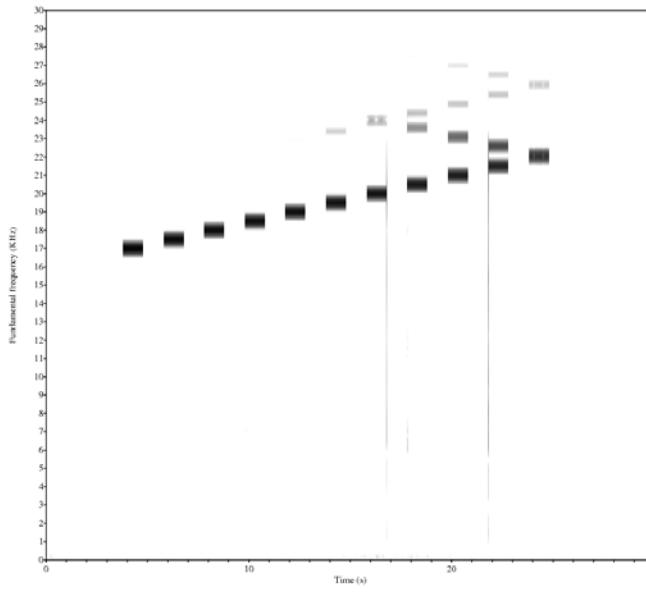
## Nokia 6210 Navigator at 60% file volume



**max volume**

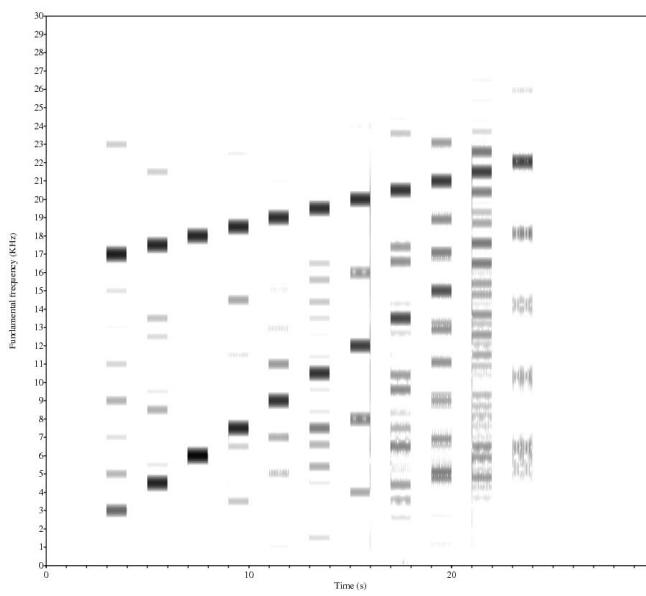


**max volume - 1**

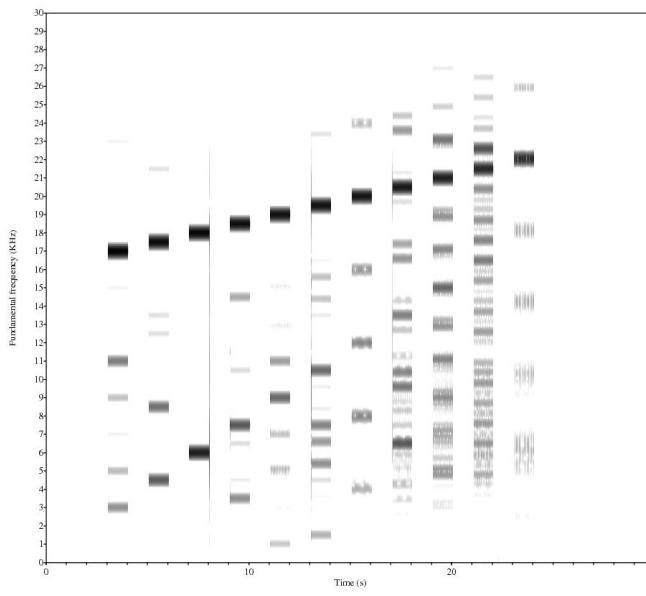


**max volume - 2**

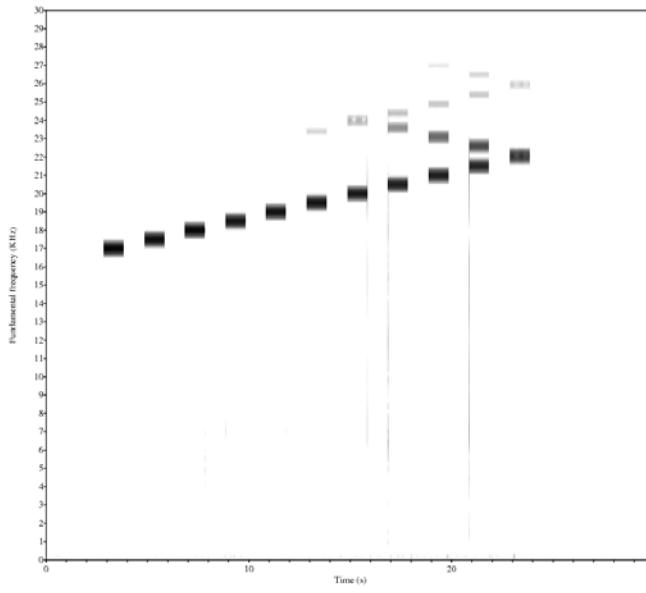
## Nokia 6210 Navigator at 80% file volume



**max volume**

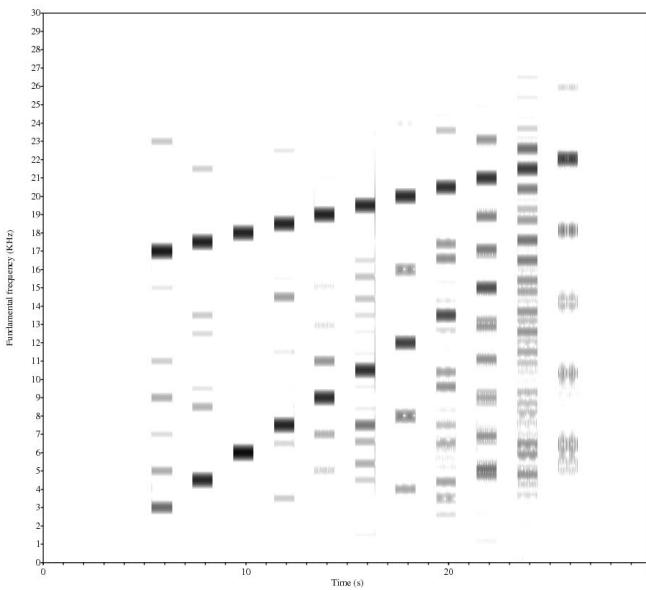


**max volume - 1**

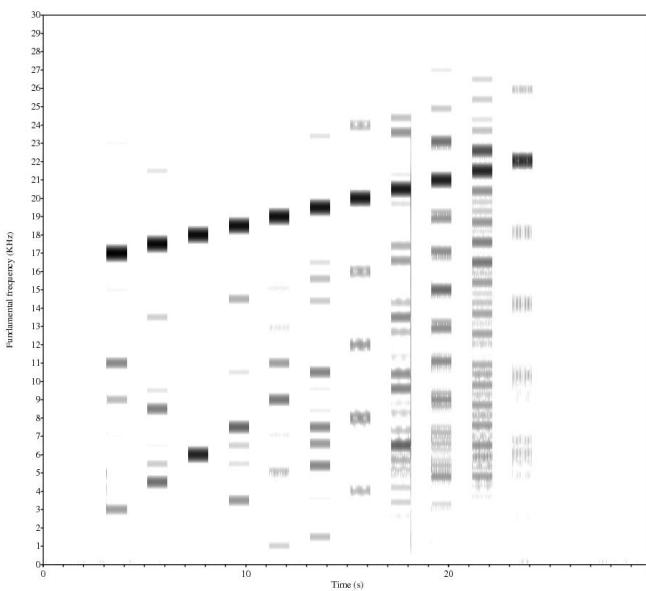


**max volume - 2**

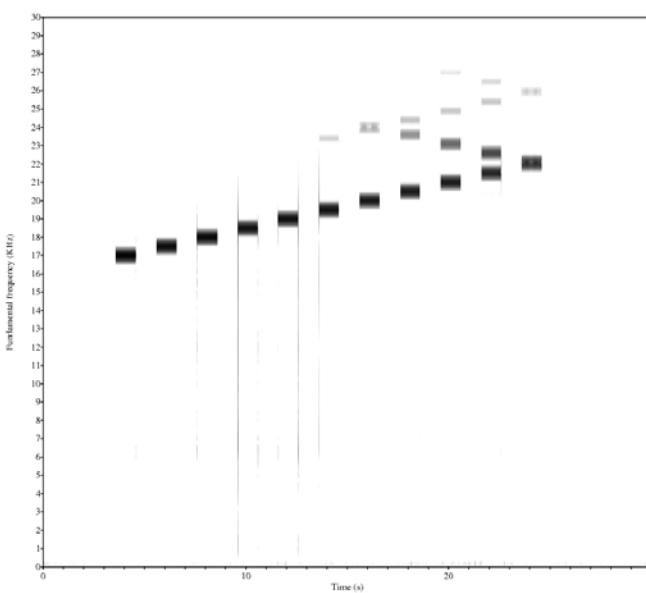
## Nokia 6210 Navigator at 100% file volume



**max volume**



**max volume - 1**



**max volume - 2**

## APPENDIX 2. RANGE MEASUREMENTS

Listed below are values collected for range experiment in Section 3.3. All values are give in dB. Columns correspond to angle between speaker and microphone (in degrees) and rows correspond to distance (in meters).

m	0	20	40	60	80	100	120	140	160	180
0.2	56.941	55.841	53.982	49.416	50.188	44.265	42.748	42.012	36.28	40.355
0.4	52.313	49.524	48.519	47.707	45.537	37.535	39.897	36.968	37.412	34.596
0.6	49.849	47.27	45.924	43.48	43.476	39.47	38.643	35.499	31.534	35.595
0.8	46.596	44.591	42.549	40.22	38.482	33.164	39.034	29.596	21.372	26.655
1	44.17	41.817	39.793	36.991	36.051	28.863	33.856	32.223	26.81	24.713
1.2	42.136	40.439	38.052	37.169	33.531	35.939	30.834	26.399	26.964	30.283
1.4	40.624	37.462	35.617	34.345	33.6	29.095	29.135	32.707	29.077	31.911
1.6	39.925	37.893	36.212	34.768	34.404	30.431	29.509	29.439	27.541	27.854
1.8	39.166	37.209	35.365	33.074	32.732	29.614	25.424	27.668	28.601	28.532
2	39.126	37.077	34.575	32.944	33.01	23.998	26.218	25.221	23.166	25.811
2.2	37.747	35.444	33.546	32.189	30.172	27.456	24.985	25.669	25.155	26.248
2.4	33.798	32.267	29.793	28.672	28.741	24.463	23.228	22.766	18.765	23.553
2.6	33.231	29.546	27.017	27.878	27.027	22.71	20.419	21.561	21.4	27.382
2.8	30.665	26.259	23.907	22.952	21.783	26.314	30.1	31.11	30.65	30.266
3	31.949	30.151	28.113	27.268	26.639	28.388	28.586	28.184	25.402	25.189
3.2	35.091	33.083	29.964	28.038	30.306	25.485	23.457	17.589	17.326	19.229
3.4	27.572	24.845	21.843	22.135	18.608	19.08	20.434	15.798	15.112	19.413
3.6	33.399	30.092	25.758	25.164	25.683	23.748	25.681	24.969	22.283	22.736
3.8	33.254	29.677	26.888	24.777	24.762	22.032	24.969	23.207	23.226	23.029
4	33.192	30.692	27.952	27.543	26.961	23.458	20.683	19.408	17.248	14.925
4.2	33.169	31.246	29.126	27.232	28.175	24.499	25.313	23.247	19.438	20.176
4.4	32.595	30.879	26.879	25.47	25.589	19.177	22.6	20.457	21.075	19.432
4.6	30.956	28.949	24.967	21.856	24.064	24.185	24.706	25.028	23.149	21.452
4.8	25.318	21.767	21.302	18.962	21.613	20.722	21.907	22.478	22.26	24.594
5	28.617	26.324	21.506	21.093	19.395	17.523	16.488	18.354	14.366	15.136
5.2	28.634	25.197	22.422	22.571	22.799	13.559	18.568	18.349	17.739	16.798
5.4	23.93	19.086	19.483	18.1	18.907	19.478	23.302	22.144	20.852	23.098
5.6	24.421	22.336	18.341	18.376	17.812	18.869	20.017	19.944	18.801	19.105
5.8	28.968	26.187	21.515	21.866	20.332	19.648	21.216	19.705	18.588	18.383
6	27.163	26.876	24.003	21.205	17.647	17.293	18.293	20.448	17.681	20.327
6.2	27.564	24.894	23.111	20.804	20.384	19.269	21.335	19.931	18.745	17.966
6.4	24.968	19.478	16.553	15.327	15.59	14.996	17.394	19.405	19.803	19.274
6.6	18.736	17.444	16.215	16.732	12.403	14.369	10.915	18.999	13.679	17.31
6.8	26.619	23.741	21.171	17.45	18.054	14.48	20.406	21.871	20.86	21.44
7	23.947	21.317	17.431	18.026	11.394	15.785	14.332	16.787	16.091	16.926
7.2	21.618	20.644	17.916	17.443	15.834	19.847	20.665	19.185	20.118	17.609
7.4	20.893	21.422	18.475	16.863	15.432	12.384	10.518	11.8	10.308	8.884
7.6	21.45	19.22	18.626	16.271	16.274	18.777	21.29	20.826	21.267	21.756
7.8	24.018	22.886	20.099	17.889	17.475	20.1	19.88	20.654	19.862	20.387
8	17.468	16.718	11.825	12.254	14.545	15.947	13.988	15.258	16.816	19.224

### APPENDIX 3. TDOA SOURCE CODE

Provided below is Java source code for TDOA trilateration.

```

// variable "measurements" contains the number of measurements

double[] x = new double[measurements]; // contains x coordinates of control points
double[] y = new double[measurements]; // contains y coordinates of control points
// first control point is expected to be
// the closest

double[] dist = new double[measurements]; // contains measurements, dist[0] is
// expected to contain 0

double[] distn = new double[measurements];

double xp=0, yp=0; // contain estimated phone coordinates

for(int i=0; i < measurements; i++)
{
    // populate x, y, and dist arrays here
}

for(int i=0; i < measurements; i++)// equation (4-10) and (4-11)
{
    xp=xp+x[i];
    yp=yp+y[i];
}

xp=xp/measurements; // make phone coordinates equal to
yp=yp/measurements; // average of all control point coordinates

int co=0; // counts the number of iterations
Matrix Xm;
double[] S = new double[unknowns];

do{
    co++;
    double[][] aMat = new double[measurements-1][2];// set the size of
    double[][] lMat = new double[measurements-1][1];// A and L matrices

    distn[0]=Math.sqrt(Math.pow(xp-x[0],2)+Math.pow(yp-y[0],2));
    // recalculate d1 using equation (4-12)

    for(int i=1; i < measurements; i++)
    {
        distn[i]=(Math.sqrt(Math.pow(xp-x[i],2)+Math.pow(yp-y[i],2)))-distn[0];
        // recalculate measurements using equations (4-13), (4-14), (4-15)

        aMat[i-1][0] = (xp-x[i])/(distn[0]+dist[i])-(xp-x[0])/distn[0];
        aMat[i-1][1] = (yp-y[i])/(distn[0]+dist[i])-(yp-y[0])/distn[0];
        // populate A matrix as in equation (4-6)

        lMat[i-1][0] = dist[i] - distn[i]; // populate L matrix as in equation (4-8)
    }

    // calculate X matrix using equation (4-5):

    Matrix A = new Matrix(aMat);
    Matrix L = new Matrix(lMat);
    Matrix At = A.transpose(); // transpose A matrix
    Matrix AtA = At.times(A); // multiply A and transposed A
    Matrix AtL = At.times(L); // multiply transposed A with L matrix
    Matrix AtAi = AtA.inverse(); // get inverse of A multiplied by A transposed
    Xm = AtAi.times(AtL); // multiply the two together
}

```

```

Matrix V = A.times (Xm);      //
V = V.minus (L);            //
Matrix VtV = V.transpose(); // calculate residuals
VtV = VtV.times (V);        //

double So = Math.sqrt(VtV.get(0, 0)/(A.getRowDimension() - A.getColumnDimension()));
// calculate standard deviation of unit weight

for(int i=0; i < unknowns; i++)
{
    S[i] = So * Math.sqrt(AtAi.get(i, i)); // each adjusted unknown
}

xp=xp+Xm.get(0, 0); // recalculate estimated phone position
yp=yp+Xm.get(1, 0);

}while (Math.abs(Xm.get(0, 0)+Xm.get(1, 0))>0.01 && co<1000);
// check if latest corrections are sufficiently low or counter reached 1000

// now "xp" and "yp" contain final estimated coordinates of the phone

// array "S" contains standard deviations

// "co" contains the number of iterations

```

## APPENDIX 4. ACCURACY EXPERIMENT VALUES

Listed below are true, average, best and worst position coordinates from experiments 1-3 in Section 5.1. The first column contains the check point number. Columns 2 and 3 contain the true X and Y coordinates of the check point. Columns 4 and 5 contain average of all readings for the given check point. Columns 6 and 7 contain the most accurate reading. Columns 8 and 9 contain the worst reading.

### **Experiment 1 (2D Trilateration)**

check point	true X	true Y	average X	average Y	best X	best Y	worst X	worst Y
<b>1</b>	1560	1960	1566	2031	1574	1969	1512	2109
<b>2</b>	1730	5490	1584	5326	1703	5451	1325	5267
<b>3</b>	2380	1015	2249	1144	2342	1031	2084	1318
<b>4</b>	2614	5570	2672	5676	2609	5572	2726	5793
<b>5</b>	3502	3497	3410	3603	3450	3527	3335	3647
<b>6</b>	4362	1820	4302	1843	4340	1819	4229	1702
<b>7</b>	4840	2783	4904	2919	4902	2840	4934	2978
<b>8</b>	5080	5483	4949	5617	5032	5446	4855	5796
<b>9</b>	5763	1635	5589	1615	5672	1626	5509	1613
<b>10</b>	5910	6193	5860	5920	5769	6142	5908	5782

### **Experiment 2 (3D Trilateration)**

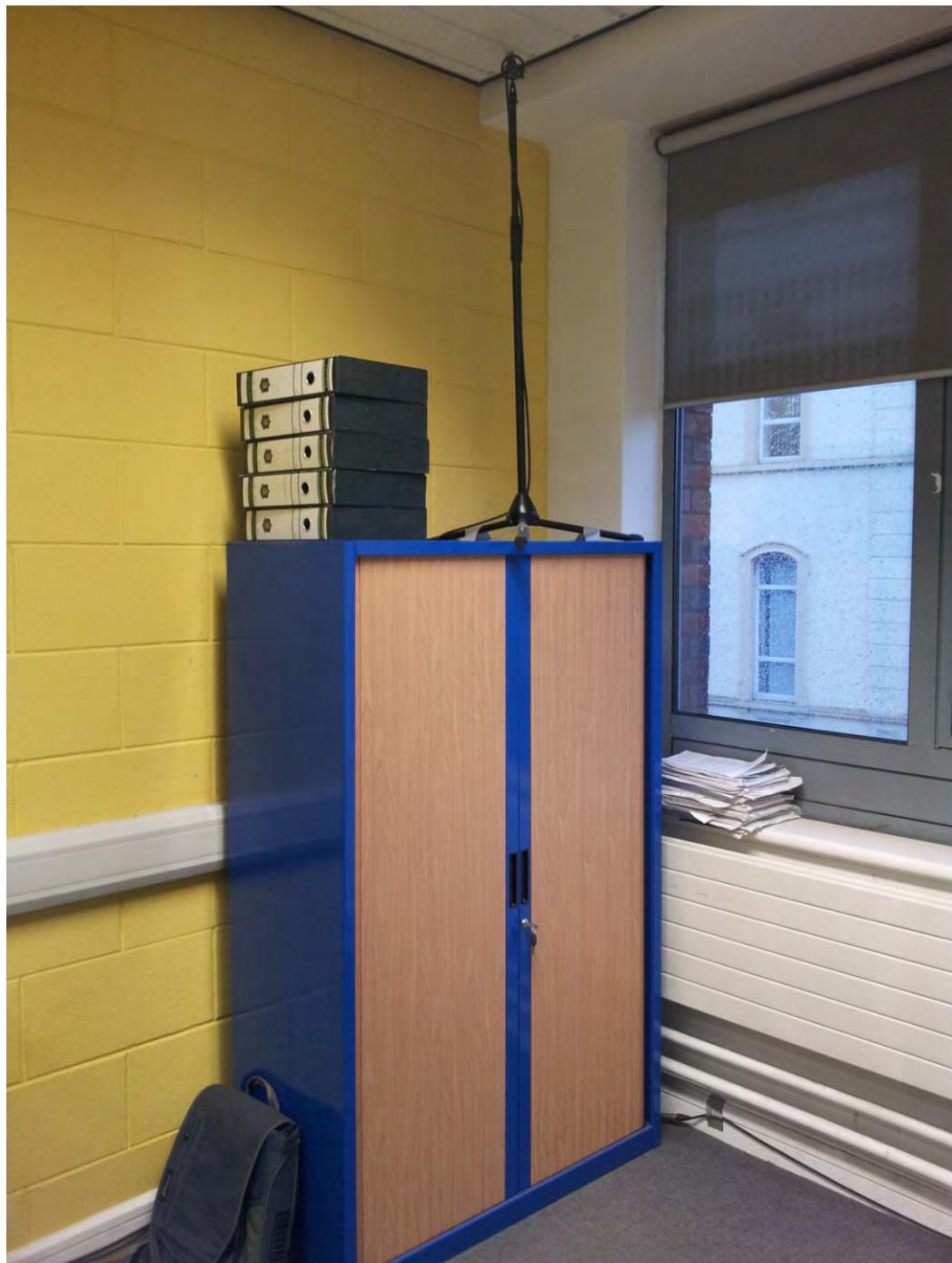
check point	true X	true Y	average X	average Y	best X	best Y	worst X	worst Y
<b>1</b>	1560	1960	1592	2129	1597	2026	1587	2202
<b>2</b>	1730	5490	1937	5604	1865	5543	1970	5782
<b>3</b>	2380	1015	2513	1235	2499	1154	2481	1368
<b>4</b>	2614	5570	2681	5468	2624	5555	2724	5373
<b>5</b>	3502	3497	3392	3591	3483	3492	3257	3624
<b>6</b>	4362	1820	4330	1983	4325	1889	4303	2060
<b>7</b>	4840	2783	4774	2842	4798	2798	4779	2929
<b>8</b>	5080	5483	4895	5262	4900	5372	4981	5142
<b>9</b>	5763	1635	5803	1813	5738	1688	5917	1888
<b>10</b>	5910	6193	5969	6052	5913	6195	5745	5652

**Experiment 3 (3D Trilateration with room calibration factor of 1.1)**

<b>check point</b>	<b>true X</b>	<b>true Y</b>	<b>average X</b>	<b>average Y</b>	<b>best X</b>	<b>best Y</b>	<b>worst X</b>	<b>worst Y</b>
<b>1</b>	1560	1960	1567	1891	1557	1962	1617	1803
<b>2</b>	1730	5490	1706	5568	1696	5508	1699	5645
<b>3</b>	2380	1015	2369	932	2377	1009	2393	796
<b>4</b>	2614	5570	2614	5656	2624	5584	2613	5724
<b>5</b>	3502	3497	3372	3466	3435	3437	3316	3478
<b>6</b>	4362	1820	4283	1812	4322	1815	4229	1824
<b>7</b>	4840	2783	4868	2690	4848	2733	4925	2654
<b>8</b>	5080	5483	5039	5528	5075	5490	5029	5647
<b>9</b>	5763	1635	5793	1622	5769	1631	5970	1780
<b>10</b>	5910	6193	5861	6108	5873	6159	5838	6047

## APPENDIX 5. EQUIPMENT

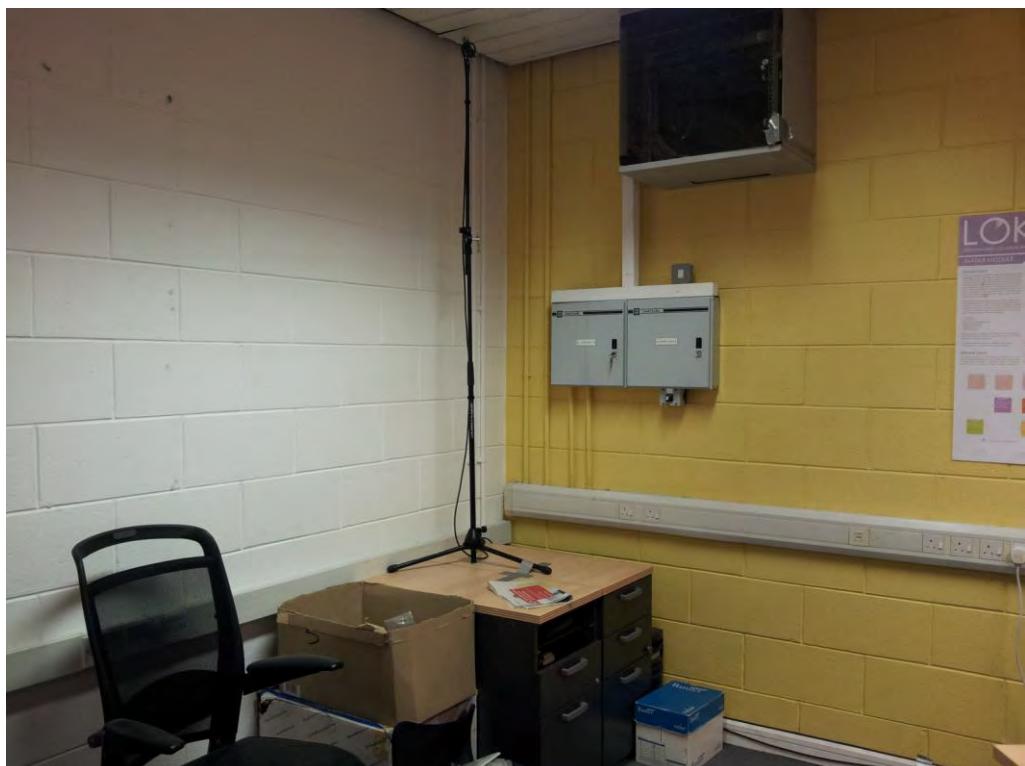
Below are pictures of equipment used in positioning experiments.



**DPA microphone 2**



**DPA microphone 1**



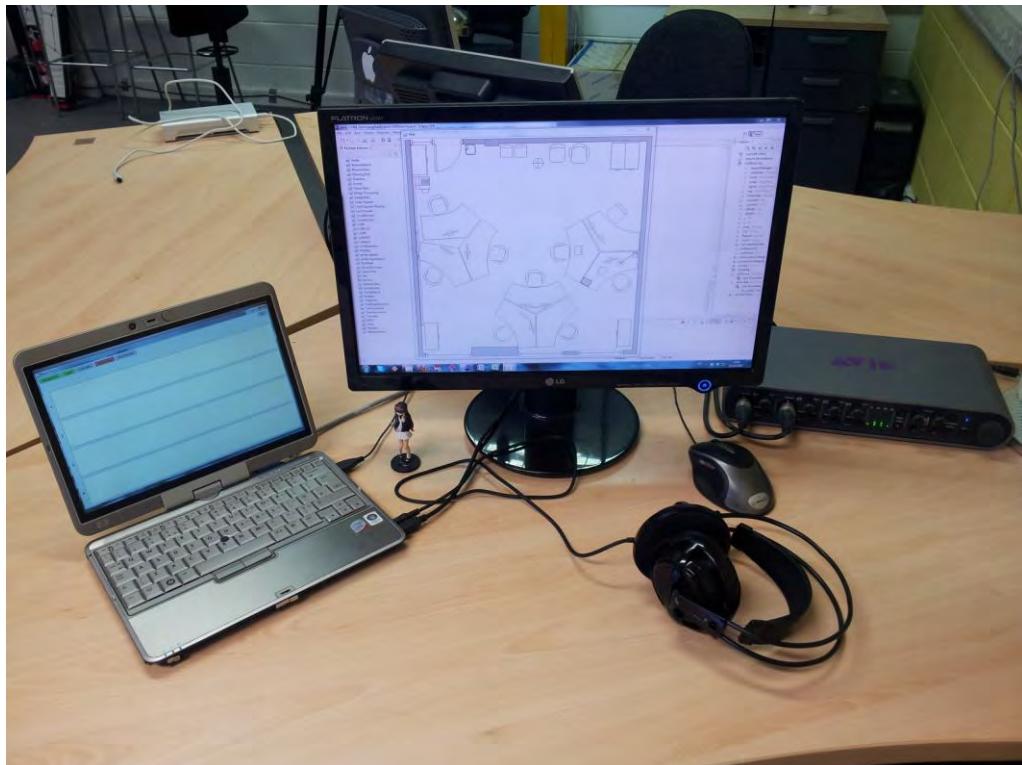
**DPA microphone 3**



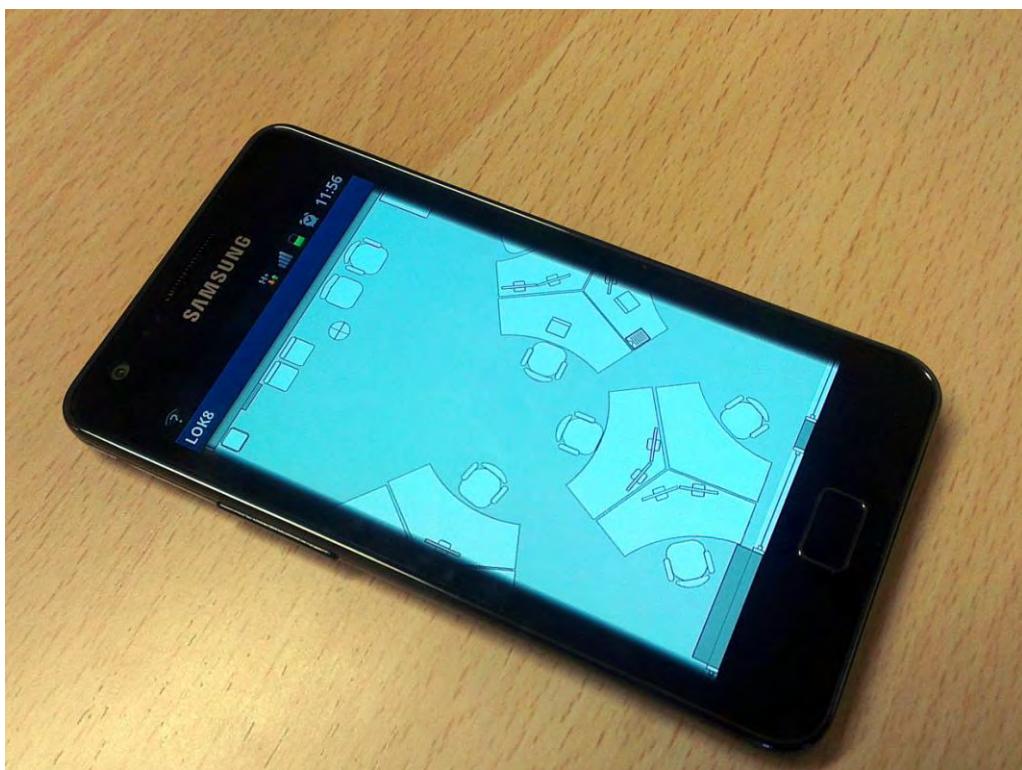
**DPA microphone 4**



**Avid Mbox Pro audiocard**



**HP Compaq 2710p laptop running LOK8 server**



**Samsung Galaxy S2 running LOK8 client**

# **Hybrid indoor positioning and directional querying on a ubiquitous mobile device**

Viacheslav Filonenko, James D. Carswell

6<sup>th</sup> International Symposium on LBS & Telecartography

2-4 September 2009,

University of Nottingham, UK

Paper 5

# Hybrid Indoor Positioning and Directional Querying on a Ubiquitous Mobile Device

Viacheslav Filonenko and James D. Carswell

*Digital Media Centre, Dublin Institute of Technology, Ireland*

{viacheslav.filonenko, jcarswell}@dit.ie

## Introduction

Spatial awareness is identified as a key feature of today's mobile devices. While outdoor navigation has been available and widely used for some time already with the help of GPS, indoor positioning has not yet made it into mainstream life. GPS and other GNSS systems offer accuracy of a scale different to that required for efficient indoor navigation. Due to this and poor signal quality in urban environments, a lot of effort has been put into developing dedicated indoor locationing systems. However, many such systems use specialized hardware to calculate accurate device position, as readily available wireless protocols have so far not delivered accuracy close to what is desired. This research aims to investigate how a number of sensors such as a Digital Compass, Bluetooth, WiFi, and Accelerometer may be combined to calculate device position and orientation to perform directional querying in a spatial database. These four technologies were chosen because they appear in some mobile devices available today and are likely to become even more widespread in the nearest future.

**Keywords:** indoor positioning, directional querying, spatial database

## LOK8 Project Overview

The LOK8 (locate) project is funded by Strand III and its goal is to create a new and innovative approach to human-computer interactions. With LOK8 a person will be able to engage in meaningful interaction with a computer interface in a much more natural and intuitive way than we are used to. A virtual character (Avatar) will be displayed in numerous locations depending on the user's position and context.

Users will be able to communicate with this virtual character through speech and gestural input/output, which will be processed and controlled by the dialog management component. This will allow "face-to-face" interactions with the LOK8 system. The LOK8 system will deliver content to the user in a variety of context-specific ways with the aim of tailoring content to suit the user's needs. In addition to screens and projectors displaying the avatar, the user's mobile device, as well as speakers within the environment, will be used to deliver focus-independent content.

Ultimately the goal is to replace a human-computer interface with a human—"virtual human" interface.

## Tracker Overview

Tracker module is one of the key components in the LOK8 system. It lets the rest of the system have access both to information about the current user's position and his surroundings. Together these make the system spatially aware.

Tracker consists of 3 components. Positioning component attempts to track the user's location throughout the program's runtime using hardware both on the phone and other parts of the LOK8 system. Environment Model stores information about the shape and size of the rooms as well as the locations and properties of objects in them. Finally Spatial Querying combines the two and allows the user to point his phone at any registered object and the premises and find out what it is.

## Positioning

### Orientation

To allow Spatial Querying the system has to be aware both of the location of the phone and its orientation. Accelerometers and the compass are primarily used for the latter.

It is possible to determine which direction a mobile phone is pointing if the following angular/spatial variables are gathered in real time: pitch angle, yaw angle and x,y,z coordinates. Pitch is an angle of rotation in the vertical plane (i.e. an angle in the up and down direction) and can be measured either from the Zenith (up) position downwards or from the Nadir (down) position upwards. (Figure 1)

Gyroscopes or accelerometers can register and present this variable. Although gyroscopes are known to be better at this task [1], they are not normally found in devices such as mobile phones and currently there is no trend that suggests that they will. Accelerometers, however, are becoming ever more popular, being used for example to automatically switch between portrait and landscape screen views on some devices currently

available today (e.g. HTC Diamond, iPhone, GPhone).

Unfortunately accelerometers can't determine yaw – rotation in the horizontal plane (i.e. an angle in left and right direction) usually measured as a compass bearing or the azimuth from North. However, yaw angle can be read from a digital compass (magnetometer). Magnetometer sensors are not yet as widely available in most modern mobile phones as are accelerometers, although it is becoming more popular of late.

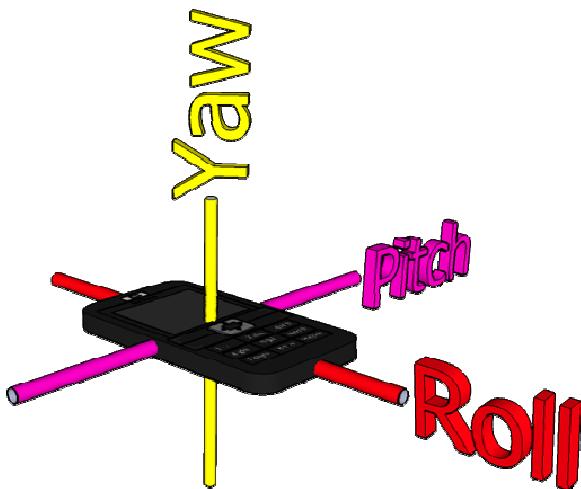


Figure 1: Roll, Yaw and Pitch axis.

### Position

Finally there are the device positional coordinates in 3D space. These three variables show where the device is located relative to a particular origin point inside the building along the x, y and z axis. These measurements can only be taken indirectly by processing Bluetooth or WiFi properties such as signal strength in some sort of trilateration adjustment. Therefore, a lot of care has to be taken into account for any unwanted interference (e.g. walls, electrical interference, reflection, etc.) that can significantly degrade the original signal strength properties [2]. Using specialized hardware could help significantly in this case, however that would seriously impede LOK8's scalability and ease of setup.

A Bluetooth beacon will be placed at the top corner of every room in the testbed environment. Other beacons will be placed in the corridors. It is proposed to implement this module as follows. (Figure 2)

1. First we determine in which room the mobile phone is right now. The easiest way to do that is to assume the user is in the same room as the closest beacon.
2. Signal strength and Bit Error rate are recorded for the other beacons in the same room. Signals from beacons that are in other rooms are easily

identified and ignored as they are greatly influenced by walls.

3. A trilateration procedure is used to calculate device position relative to the known positions of fixed beacons. It may also be useful to take differences between ceiling height and a device's position into account.
4. The local position in the room is then converted to the relative position in the premises and may be further converted to absolute coordinates in real-world space if required for seamless indoor/outdoor navigation and wayfinding.
5. Parallel to Bluetooth positioning, accelerometers will work in both movement and rotation modes to track a user's movements. If successful this technique will be similar to dead reckoning, and can be used in a number of ways. First of all the program can generate a path the user has walked so far. When a user enters a room, the path can be checked against the layout of known obstacles stored in the database and help correct the user's current position. Also it can be used to determine which of the signals is blocked by the user and accordingly apply appropriate weights in the trilateration procedure. [3]

### Environment Model

There will be a central spatial database accessible through Bluetooth. There will be an entry in the database for each beacon's ID, xyz position, and distance to other beacons in the same room, along with the room ID. At some point, attributes of objects (e.g. desks, posters, paintings) will be added to the dataset as well. These various objects will carry position, dimensions and description attributes (e.g. whose desk it is, what poster is it, whose office is it).

### Spatial Querying

After an accurate position and orientation have been determined, it is possible to find out which object, if any, the phone is pointing at. This will only be done when the user presses a button associated with querying. We will assume that the phone in this case is used in the same way as a television remote control – e.g., the top end of the phone points in the direction of the object of interest. Once the query parameters have been captured and the query processed, the phone will beep to let the user know a query result has been returned to the screen. If no object was identified a doublebeep will sound.

Identifying an object in the room could be done through ray-box collision detection in 3D space. This can be achieved either externally using existing ray-box collision detection algorithms or inside the spatial database itself, if it supports such ray intersection queries in 3D. [4]

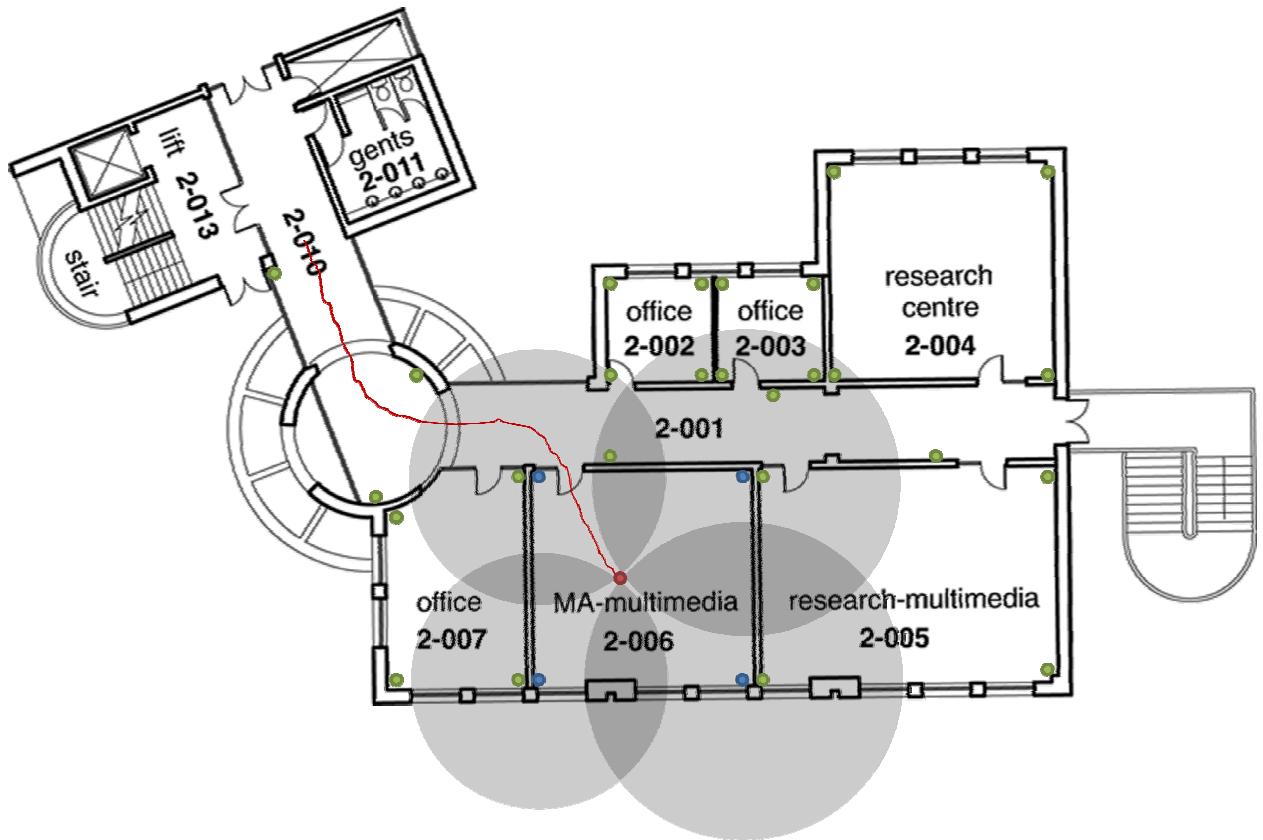


Figure 2: The red dot is user's location. Blue dots are beacons that are currently used for positioning. Green dots are other beacons. Red line is user's route as traced by the system.

## References

- [1] Invensense. Gyroscopes and Accelerometers compared Video.  
[http://invensense.com/support/FLVPlayer\\_Progressive.swf?skinName=Halo\\_Skin\\_3&streamName=../Library/InvenSeInv5\\_VP6\\_512K](http://invensense.com/support/FLVPlayer_Progressive.swf?skinName=Halo_Skin_3&streamName=../Library/InvenSeInv5_VP6_512K)
- [2] Kolodziej, K., Hjelm, J., Local Positioning Systems, Taylor & Francis Group, 2006.
- [3] Hallberg, J., Nilsson, M., Synnes, K. "Positioning with Bluetooth.", 10th International Conference on Telecommunications, ICT, 2003
- [4] Williams, A., Barrus, S., Morley, K., Shirley, P. "An efficient and robust Ray-Box intersection algorithm.", Journal of Graphics Tools, A K Peters, pp. 49-54, 2005

# **Tracker: indoor positioning for the LOK8 project**

Viacheslav Filonenko, James D. Carswell

9<sup>th</sup> IT & T Conference

22-23 October 2009,

Dublin Institute of Technology, Ireland

Paper 25

# **Tracker: Indoor Positioning for the LOK8 Project**

**Viacheslav Filonenko, James D. Carswell**

Digital Media Centre, Dublin Institute of Technology, Ireland  
{viacheslav.filonenko, jcarswell}@dit.ie

## **Abstract**

Spatial awareness is identified as a key feature of today's mobile devices. While outdoor navigation has been accessible and broadly used for some time already with the help of GPS, indoor positioning has not yet made it into mainstream life. GPS and other GNSS systems offer accuracy of a scale different to that required for efficient indoor navigation. This research aims to investigate how a number of sensors such as a Digital Compass, Bluetooth and Accelerometer may be combined to calculate device position and orientation to perform directional querying in a spatial database. These three technologies were chosen because they appear in some mobile devices available today and are likely to become even more widespread in the nearest future.

**Keywords:** indoor positioning, directional querying, location based services

## **1 LOK8 Project Overview**

The LOK8 (locate) project is funded by Strand III and its goal is to create a new and innovative approach to human-computer interactions. With LOK8 a person will be able to engage in meaningful interaction with a computer interface in a much more natural and intuitive way than we are used to. A virtual character (Avatar) will be displayed in numerous locations depending on the user's position and context. Users will be able to communicate with this virtual character through speech and gestural input/output, which will be processed and controlled by the dialog management component. This will allow "face-to-face" interactions with the LOK8 system. The LOK8 system will deliver content to the user in a variety of context-specific ways with the aim of tailoring content to suit the user's needs. In addition to screens and projectors displaying the avatar, the user's mobile device, as well as speakers within the environment, will be used to deliver focus-independent content. Ultimately the goal is to replace a human-computer interface with a human-“virtual human” interface.

## **2 Tracker Overview**

Tracker module is one of the key components in the LOK8 system. It lets the rest of the system have access both to information about the current user's position and his surroundings. Together these make the system spatially aware. Tracker consists of 3 components. Positioning component attempts to track the user's location throughout the program's runtime using hardware both on the phone and other parts of the LOK8 system. Environment Model stores information about the shape and size of the rooms as well as the locations and properties of objects in them. Finally Spatial Querying combines the two and allows the user to point his phone at any registered object and the premises and find out what it is. This poster for the 9th IT&T conference summarizes the work presented at 6th International Symposium on LBS & TeleCartography [1].

### 3 Related Work

There are a number of locationing services that operate on a larger (outdoor) scale. First of all there is GPS and GLONASS, which make use of trilaterating signals transmitted from satellites. These don't work indoors very well and the average accuracy is found to be in the neighbourhood of 15 meters in urban environments [2]. Then there's Assisted GPS (A-GPS) which improves the startup "fix" time and accuracy in urban environments by accessing some rough positioning of visible satellite information (ephemeris data) through GPRS. Cell tower triangulation is also an emerging service. Its reported accuracy however is between 50 and 300 meters depending on atmospheric conditions and tower dispersion geometry [3].

Another approach is to read MAC addresses and associated signal strengths of all currently accessible WiFi access points and calculate position through trilateration. This service is currently offered commercially by Navizon and Skyhook bundled with cell tower triangulation and optionally GPS [3,4]. Their services are designed to either replace GPS, for example on mobile devices without GPS receivers, or enhance its accuracy in urban environments. However the resulting accuracy is still roughly in the 10-20 meter range.

mong locationing systems currently published, there are some that achieve a much higher level of accuracy using specialised client-side hardware. The Bat and The Cricket both use ultrasound, for example, to measure distance to receivers placed on the ceiling in a grid, but do so in different ways [5,6]. In case of the Bat transmitter, the device transmits a short ultrasound pulse, the time-of-flight from the transmitter to receivers mounted at known positions is measured. Because the speed of sound in the air is known, distance to each of these transmitters can be calculated and then used to calculate the exact position of the transmitting device using trilateration.

### 4 Positioning

To allow Spatial Querying the system has to be aware both of the location of the phone and it's orientation.

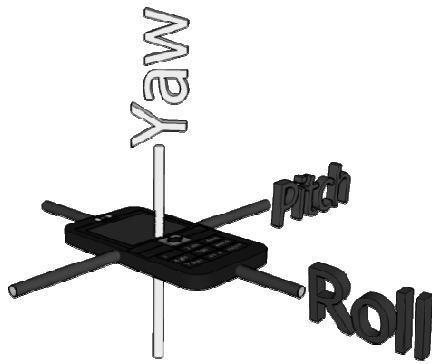
It is possible to determine which direction a mobile phone is pointing if the following angular/spatial variables are gathered in real time: pitch angle, yaw angle and x,y,z coordinates. Pitch is an angle of rotation in the vertical plane (i.e. an angle in the up and down direction) and can be measured either from the Zenith (up) position downwards or from the Nadir (down) position upwards. (Figure 1)

Two of the three variables can be registered by accelerometers, which are becoming ever more present in modern mobile devices. Unfortunately accelerometers can't determine yaw – rotation in the horizontal plane (i.e. an angle in left and right direction) usually measured as a compass bearing or the azimuth from North. However, yaw angle can be read from a digital compass.

The device's position in space will be detemined through trilateration. A lot of care has to be taken into account for any unwanted interference (e.g. walls, electrical interference, reflection, etc.) that can significantly degrade the original signal strength properties [7]. A Bluetooth beacon will be placed at the top corner of every room in the testbed environment. Other beacons will be placed in the corridors. It is proposed to implement this module as follows. (Figure 2)

1. First we determine in which room the mobile phone is right now. The easiest way to do that is to assume the user is in the same room as the closest beacon.
2. Signal strength and Bit Error rate are recorded for the other beacons in the same room. Signals from beacons that are in other rooms are easily identified and ignored as they are greatly influenced by walls.
3. A trilateration procedure is used to calculate device position relative to the known positions of fixed beacons.

4. The local position in the room is then converted to the relative position in the premises.
5. Parallel to Bluetooth positioning, accelerometers will work in both movement and rotation modes to track a user's movements. If successful this technique will be similar to dead reckoning, and can be used in a number of ways.



**Figure 1: Roll, Yaw and Pitch axis.**

## 5 Environment Model

There will be a central spatial database accessible through Bluetooth. There will be an entry in the database for each beacon's ID, xyz position, and distance to other beacons in the same room, along with the room ID. At some point, attributes of objects (e.g. desks, posters, paintings) will be added to the dataset as well. These various objects will carry position, dimensions and description attributes (e.g. whose desk it is, what poster is it, whose office is it).

## 6 Spatial Querying

After an accurate position and orientation have been determined, it is possible to find out which object, if any, the phone is pointing at. This will only be done when the user presses a button associated with querying. We will assume that the phone in this case is used in the same way as a television remote control – e.g., the top end of the phone points in the direction of the object of interest. Once the query parameters have been captured and the query processed, the phone will beep to let the user know a query result has been returned to the screen. If no object was identified a doublebeep will sound.

Identifying an object in the room could be done through ray-box collision detection in 3D space. This can be achieved either externally using existing ray-box collision detection algorithms or inside the spatial database itself, if it supports such ray intersection queries in 3D [8].

## 7 Current Work

Currently all four LOK8 modules are collaborating on setting up the “Wizard of Oz” test environment. This should let us simulate the surface functions of the system and then record and analyze user's interaction with it. The results of analysis will be used to improve the interface and see how the setup makes the user experience different, or what improvements/drawbacks it presents. In respect to the Tracker module it will be useful to see what level of accuracy the user expects or can tolerate, as well as how exactly the user makes queries or does other interactions using the phone. A detailed description of the setup can be found in [9].

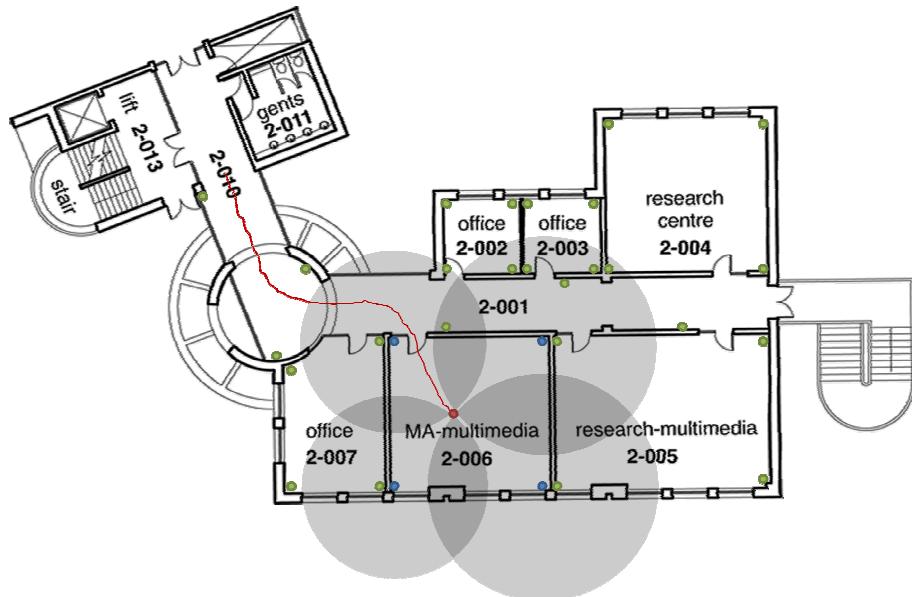
## 8 Conclusions

The focus of our upcoming work therefore involves experimenting with the Bluetooth, magnetometer, and accelerometer sensors on the phone. Sensor fusion research into finding the most responsive and

efficient combination of these three different sources of information about a device's movement will be made, followed by creating a reliable locationing framework on which to build our Lok8 spatial query system. One idea is to use beacons in a different way by limiting their transmitting area with a form of Faraday cage. If beacons are positioned in such a way that their areas overlap, this will prove to be a reliable source of information as certain beacons can only be detected within certain areas of the room. Combining this with information gathered by the accelerometers, it may then be possible to achieve the higher locational accuracy we require for accurate indoor cellphone positioning for targeted 3D directional querying.

## 9 Acknowledgements

The authors wish to thank the Higher Education Authority (HEA) in Ireland and specifically their Technological Sector Research Strand III: Core Research Strengths Enhancement Programme for funding the work carried out at the Dublin Institute of Technology on the Lok8 project.



**Figure 2: Bluetooth beacon layout for the LOK8 environment.**

## References

- [1] Filonenko, V., Carswell, J. "Hybrid Indoor Positioning and Directional Querying on a Ubiquitous Mobile Device", published in proceedings of 6th International Symposium on LBS & TeleCartography, Nottingham, UK, 2009
- [2] Modsching, M., Kramer, R., Klaus, H. "Field trial on GPS Accuracy in a medium size city: The influence of built-up.",
- [3] Navizon Technical Paper, 2007.  
[http://www.navizon.com/Navizon\\_wifi\\_gps\\_and\\_cell\\_tower\\_positioning.pdf](http://www.navizon.com/Navizon_wifi_gps_and_cell_tower_positioning.pdf)
- [4] Skyhook in Action. <http://www.skyhookwireless.com/inaction/>
- [5] The Bat Ultrasonic Location System Website.  
<http://www.cl.cam.ac.uk/research/dtg/attarchive/bat/>
- [6] Cricket v2 User Manual 2005. <http://cricket.csail.mit.edu/v2man.pdf>
- [7] Kolodziej, K., Hjelm, J., Local Positioning Systems, Taylor & Francis Group, 2006.
- [8] Williams, A., Barrus, S., Morley, K., Shirley, P. "An efficient and robust Ray-Box intersection algorithm.", Journal of Graphics Tools, A K Peters, pp. 49-54, 2005
- [9] Schütte, N., Kelleher, J., Mac Namee, B. "A Mobile Multimodal Dialogue System for Location Based services", awaiting publication in IT&T 2009 proceedings

# **Investigating Ultrasonic Positioning on Mobile Phones**

Viacheslav Filonenko, Charlie Cullen and James Carswell

International Conference on Indoor Positioning and Indoor Navigation

(IPIN 2010)

15 – 17 September 2010,

ETH Zurich, Switzerland

Pages 419-426

IEEE Xplore 2010

# Investigating Ultrasonic Positioning on Mobile Phones

Viacheslav Filonenko, Charlie Cullen and James Carswell

Digital Media Centre, Dublin Institute of Technology, Ireland

{viacheslav.filonenko, charlie.cullen, jcarswell}@dit.ie

**Abstract**—In this paper we evaluate the innate ability of mobile phone speakers to produce ultrasound and the possible uses of this ability for accurate indoor positioning. The frequencies in question are a range between 20 and 22 KHz, which is high enough to be inaudible but low enough to be generated by standard sound hardware. A range of tones is generated at different volume settings on several popular modern mobile phones with the aim of finding points of failure. Our results indicate that it is possible to generate the given range of frequencies without significant distortions, provided the signal volume is not excessively high. This is preceded by the discussion of why such ability on off-the-shelf mobile devices is important for Location Based Services (LBS) applications research. Specifically, this ability could be used for indoor sound trilateration positioning. Such an approach is uniquely characterized by the high accuracy inherent to sound trilateration, with little computational burden on the mobile device, and no specialized hardware or audible noise. Combined with a fast internet connection and other sensors present in modern smartphones, such as accelerometers and magnetometers, our approach confirms mobile phones as a suitable platform for indoor LBS applications.

**Keywords**—Ultrasound; Indoor Positioning; Mobile Devices

## I. INTRODUCTION

Currently outdoor Location Based Services (LBS) have the advantage of reliable positioning via GPS (also Wi-Fi and GSM) and a defined business model for the delivery of content to the user. This has led outdoor LBS to greatly expand in recent years, though indoor locationing technologies and methods have yet to fully mature on mobile devices. In the current state of the art for indoor LBS, merging accurate (i.e. sub-metre) indoor positioning and context-sensitive services is still an outstanding problem. Existing systems such as employee tracking [1] using RFID/Wi-Fi tags or badges are relatively cheap to implement, but no development path for mobile device RFID currently exists in Europe. For context-sensitive services, such as a virtual tour guide, factors such as device cost, functionality and service provision are still stumbling blocks to effective implementation of solutions. A frequent example would require the user to point a device at a tag or enter an exhibit's number manually. Such approaches are time consuming, complex and require user focus (thus distracting them from the exhibits). In addition, inability to provide effective user navigation (e.g. how to find an exit) and lack of rich media multimodal interfaces has led to a disparity between device capabilities (where media delivery is a de facto standard) and quality user focused services.

Currently there are no examples of fully-functional indoor LBS for mobile phones, but theoretically they could perform a number of functions:

- Make evacuation procedure more intuitive and efficient by showing directions along the shortest path [2]. In this example it is important for the system to know 100% of the time where the user is so that they do not have a reason to panic if suddenly realising that they are lost.
- Improve navigation in shopping malls. There is already a company that collects and maintains maps of shopping malls [3]. Normally when working with an unfamiliar map it takes a significant amount of time to figure out current position and direction unless the map is stationary and the position is already marked. This makes portable maps less useful. Using indoor positioning it is possible to take better advantage of such data. Showing the current position on an interactive map would already be a significant improvement and giving instructions how to get to a particular shop would make navigation easier still.
- Given better accuracy, it may be possible to direct the user to a particular shelf in a shop. Bearing that in mind it is possible to design a program where the user has populated on their mobile phone a list of things they need to buy since they last went shopping. When they enter a shop, the most optimal route to collect the goods is generated and the user is instructed where to go next.
- A library catalogue combined with a navigation system that directs the user to the shelf with the book he requested.
- A museum virtual tour guide. Systems currently used in museums provide unsophisticated functionality which is very often limited to pointing at a tag or manually entering a number in order to hear a recording. A system with true indoor positioning based on a mobile phone can be used by pointing at the actual exhibit via directional querying. Depending on the arrangement and size of exhibits, directional querying may require very high spatial and directional accuracy. A smartphone can deliver a variety of content including audio, video, text, images or a combination of them such as a webpage. Once again because the system is continuously aware of user's location and orientation it is possible to guide the user to an exhibit he wants to see, to the exit, or any other facility within the museum.
- Used by a company to track employees. Systems currently used for this purpose use Wi-Fi or RFID tags. The main problem with using tags is that while the person controlling the system knows where everyone is, an average user has no benefit from this system. A smartphone version however can allow any employee to find any other employee regardless of where they are right

now. Depending on the type of work this ability may turn out to be extremely valuable. Also it is not unusual for companies to issue smartphones such as Blackberries to every employee, so it is very likely that everyone is already carrying the necessary hardware.

Section 2 of this paper discusses related work. Section 3 discusses our methodology and Section 4 presents results of our experiments. Finally Section 5 concludes the paper and presents directions of future work.

## II. RELATED WORK

### A. Indoor Positioning

Positioning on mobile phones is not limited to GPS. Other sensor components commonly found in mobile phones can also be used to determine position. Methods that use propagation of Radio Frequency (RF) signals are prevalent in this field, with the exception of computer vision, where SLAM appears to be the most promising but considered by many an operational technology still in its infancy [4]. Computer vision, although often very accurate, is characterized by high computational load, complicated procedures of recovery from tracking failures and susceptibility to camera shake and motion blur. These problems are addressed in the studies done by Williams et al. [5] and Wagner et al. [6]. Another difficulty associated with computer vision is that the user is supposed to be looking through the display screen when using the device.

Every modern smartphone at least has GSM, Wi-Fi and Bluetooth modules. Five meter accuracy, one of the best results for indoor GSM positioning, was displayed by Otsason et al. with the help of wide signal-strength fingerprinting [7]. Unfortunately wide signal-strength fingerprinting is impossible on many modern phones due to OS restrictions. Other GSM positioning methods are generally impractical for indoor use due to poor accuracy. Wi-Fi positioning on average shows twice as better accuracy than GSM. A method proposed by Ferris et al. where Gaussian processes are used to mathematically predict signal strength in areas outside the exact spots where fingerprints were taken appears to be promising [8]. The best accuracy among commercial solutions was shown by Ekahau: 1-3 meters [1]. Because of the ability to leverage hardware already present in office areas Wi-Fi is a good choice for positioning, but it will become even better when client-to-client connections are possible with Wi-Fi Direct, which is due to appear in 2010 [9]. Bluetooth has the shortest range among the three technologies. There are two major problems that make Bluetooth positioning particularly difficult. First of all it is designed to adjust signal strength when signals become too strong or too weak. Disabling this feedback loop is discussed by Zhou et al. [10]. Another problem is that it takes a lot of time for a new device to be fully discovered. Very often it means that the user has already left the area [11]. This makes Bluetooth trilateration impractical; however coarser room-level positioning can be done relatively quickly as device pairing is not required.

Currently it is impossible to achieve accuracy below one meter [12] using RF-based technologies present in mobile phones [7, 8, 13]. Time-of-arrival does offer robust performance [11], however for RF this requires specialised equipment, which is why less direct approaches using signal

strength and bit error rate have to be used. Sound, being significantly slower than RF, is easily localised to a few centimetres (due to longer time of arrival). Borriello et al. [14] showed that it is possible to emit 21 KHz (just above the human hearing range) signal from a mobile phone speaker and successfully receive with a conventional microphone. In a separate study Peng et al. [15] showed that it is possible to utilize sound in order to measure the distance between two mobile phones using time-of-arrival. These two principles are combined in our method that involves trilateration of an inaudible ultrasound signal using a static microphone array. Sound positioning is discussed in greater detail in the next section.

The comparison of positioning methods available for most smartphones is given in Table 1.

TABLE I.  
COMPARISON OF POSITIONING METHODS FOR SMARTPHONES.

	works indoor	accuracy	infrastructure cost	reliability
GPS	no	poor (n/a)	none	good
GSM	yes	average	none	good
Wi-Fi	yes	good	none/average	good
Bluetooth	yes	good	average	poor
Sound	yes	excellent	average/ expensive	good
Computer Vision	yes	excellent	none-average	poor

### B. Sound Positioning

Sound is a mechanical wave which travels at speeds much lower than the speed of light. In dry air at a temperature of 25°C the speed of sound is only 346 m/s. At such propagation speeds, one sample of a standard 44.1 KHz stream (44100 cycles/second) accounts for 0.8cm [7, 16]. In other words a signal will travel only 0.8 centimeters in the duration of the smallest time grain. Technically it is possible to work with sound even at 384 KHz, which can give much finer accuracy. Unfortunately, an audio recording does not have a reference point for when the signal was sent, it has to be collected therefore from the sender. If the sender and receiver have clock skew/drift between each other, this will result in synchronization uncertainty. One more uncertainty results from possible misalignment between the time a command to emit sound was issued and the actual emission time. Finally, receiving uncertainty occurs as a possible delay in the signal being promptly recognised.

Peng et al. showed that all of the above uncertainties can be eliminated when estimating distance between two devices [8]. Their “BeepBeep” ranging procedure involves two mobile devices starting to record sound before emitting short sound signals one after another. This way each recording has two reference points. Device A has a recording of the signal emitted by Device A reaching the microphone on Device A, and later of the signal emitted by device B reaching device A. Device B has a recording of the signal from Device A reaching Device B followed by the signal from Device B reaching Device B. The span between the two signals on Device A is longer than on device B since Device A was the

first one to emit sound. When the second span is subtracted from the first span the result is equal to twice the time it takes sound to travel between the two devices. (Figure 1)

“BeepBeep” has presented itself very well in open environments, but unfortunately showed poor accuracy indoors at distances longer than 5 meters. Most likely this was caused by the multipath effect. The experiments were done in a small room with one or the other device close to a wall, which interprets a signal that bounced off a wall to be of comparable strength to one that arrived via the shortest path.



Figure 1. BeepBeep signal exchange. The two horizontal lines represent recordings on each of the devices. Black boxes are actual sound signals that were recorded. The dashed lines represent events in time. Time interval between the two boxes on recording A minus time interval between the two boxes on recording B equals  $2 \times$  the time it takes for the signal to travel between the two devices.

“BeepBeep” presents a very good idea that overcomes several problems common to acoustic ranging systems, but unfortunately the procedure is not very suitable for locationing via trilateration. To provide the necessary measurements, there has to be at least three or four visible beacons which allows for measuring distance to them simultaneously either by listening to sound signals emitted by the mobile device or simultaneously emitting sound. The first approach seems to be intuitively favourable. Although it does not really eliminate any synchronization problems, many difficulties can be avoided by listening to just one signal at multiple locations. First of all, there is no need to distinguish between several different signals that arrive either simultaneously or very close to each other. Secondly, the computational load of trilateration will be on the server connected to the microphones, rather than the mobile device.

The effective range of transmitting beacons greatly depends on the volume of the signal and the direction of the speaker. Traditionally, a spherical model is used for sound propagation. However, it has also been observed that ultrasound fading follows a water-drop shaped model as in Figure 2, which should be true for sound at higher audible frequencies as well [7, 8]. Another thing to take into account is the fact that sound at higher frequencies can be easily blocked by furniture. Most smartphones have both a speaker and a microphone on the same side as the display screen while some also have a louder speaker on the opposite side. Regardless if the phone emits or listens for signals, beacons placed on the ceiling will have a direct line of sight with the phone’s speaker/microphone while the user is using the phone. For small rooms it should be enough therefore to place a beacon at the top of every corner of the room. Unfortunately the water-drop model suggests that if a room is significantly larger, the angle between a speaker and a microphone will be too great and the signal will fade too

much, in which case a number of beacons will have to be placed on the ceiling to form a grid. We suspect that placing microphones flat against walls/ceiling should effectively counter the multipath effect, which speaks in favour of a mobile phone as the signal source.

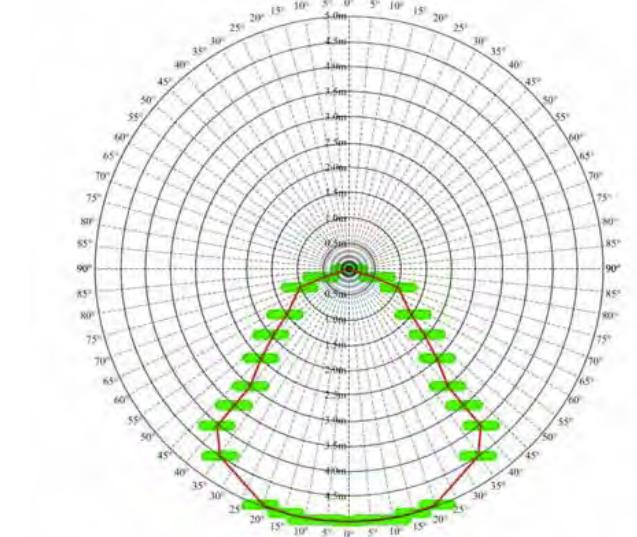


Figure 2. Directional model sound transmission, adapted from Hsiao C.[17].

It is evident from examples given above that the mobile device needs to communicate with the infrastructure somehow, first to communicate the intention to estimate position and secondly to exchange measurement results. It appears impossible to reliably transfer data with conventional speakers and microphones. According to research, the signal to noise ratio even at a range as short as 1 meter is too high to correctly decode more than 95% of the packets [7]. Wi-Fi communication is a more reliable alternative. As a result the sound signal can be of any length, shape and frequency as long as it can be reliably detected. It has been observed that the first few milliseconds of a sample playback come with a very large distortion which at certain frequencies appear to be a loud unpleasant click [7, 18]. It is therefore recommended to linearly increase the amplitude of the signal. Regrettably, this may introduce some uncertainty to where the beginning of the signal is - an otherwise perfect candidate for a reference point. The end of the signal is unsuitable because it is likely to merge with an echo coming by an alternative path. The multipath effect is also the reason why it is not efficient to determine the middle of the signal and use that as a reference. The best solution appears to be a signal that linearly increases in amplitude and immediately decreases. This will form a “peak” that the receiver will try to detect. Finally the sound frequency presents a choice between efficiency and usability. It has been suggested that anything above 8 kHz attenuates too quickly. On the other hand it appears desirable to use a frequency that is inaudible to the human ear. Frequencies above 20 KHz (ultrasound) generally cannot be picked up by human hearing. While these frequencies reduce the effective range of our system, this is offset by a noiseless positioning system placing more importance on user experience. If necessary, this would justify an increase in the number of necessary beacons. Also

higher frequencies are easily stopped by obstacles, while lower frequencies can even penetrate walls. If taken into account when designing the system either could be used to an advantage.

There are two well known examples in the literature of indoor positioning systems that successfully utilise ultrasound signals: the Bat and the Cricket. These two systems are very similar as both require a dense grid of sensors on the ceiling. Both the beacons and positioning devices are specialised hardware, designed to operate in the ultrasound range. In case of the Bat transmitter, the mobile device transmits a short ultrasound pulse and the time-of-flight from the transmitter to receivers mounted at known positions is measured. Cricket on the other hand uses a combination of radio signals and ultrasound. Beacons periodically transmit “advertisements” on a radio-frequency channel and send an ultrasonic pulse at the same time. Once the locationing device detects an advertisement, it listens for the corresponding ultrasonic pulse. Once the pulse is received, it is possible to calculate the distance by comparing the arrival time of radio and ultrasound signals. Both systems have accuracy of about 3cm, with the Cricket being slightly more accurate. Also both systems have proved to be highly scalable, being able to operate on multiple devices and over large areas. For example the Bat system was installed throughout a three-floor 10,000 square foot office building with 750 beacons, and continuously tracked 200 mobile devices [12].

It was shown by Borriello et al. that 21 KHz signals can be successfully emitted and received with conventional desktop speakers and microphones (on a HP iPAQ 3870 PDA and a Dell Inspiron 8200 laptop) [14]. The signal was also successfully detected 100% of the time within a range of 10 meters. This was done using three instances of the Goertzel algorithm: one in the 21 KHz frequency and the other two in adjacent frequencies above and below. The first instance was checked against the other two in order to distinguish the signal from background noise. In order to check how well the detection system copes with common environmental noise three separate tests were performed. One involved a number of people having a conversation, the second involved playing a variety of music recorded in two different formats (mp3 and ogg), and the final test was leaving the system running in an office environment for two consecutive days. During the three tests the detection algorithm did not detect any signals. This is a very encouraging finding, because it means that it may be possible keep working with “raw” sound without introducing complicated filters to check for false positives. The only source of false signals remains the multipath effect, which we hope can be countered with correct placement of microphones and some adjustments in detection algorithms like those proposed in [15].

Overall the ultrasound approach is an ideal solution for indoor positioning in terms of accuracy. It easily passes the one-meter threshold and comes very close to the one centimetre threshold. So far it has been implemented and tested with the help of custom hardware, but we see no reason why it could not be done using conventional speakers and microphones. Our current research therefore focuses on finding a way to implement it on conventional off-the-shelf hardware, potentially making it very cheap and accessible as

both microphones and speakers are mass produced and widely available.

### III. PROPOSED SOLUTION

After having reviewed positioning methods available on most modern smartphones, ultrasound trilateration was recognized as a suitable method to deliver fine-grained indoor positioning for the following reasons:

1. Among the positioning methods reviewed, only sound positioning can potentially offer consistent sub-meter accuracy. There are good reasons to aim for higher accuracy of estimated position and orientation. To begin with, everything indoors happens on a smaller scale. Corridors are narrower than streets and room entrances are smaller than shop fronts. An indoor LBS is very easy to expand in terms of functionality once all the infrastructure and spatial data is there, so if there is no need for sub-meter accuracy initially, lack of it should not be a limiting factor for expansion. The requirements for accuracy can be different depending on the task. For example a virtual tour guide with spatial querying will require as fine accuracy as possible, at least below one meter, because direction deviation will increase as the distance to the object increases. While privacy is a good reason to limit maximum positioning accuracy for pervasive technologies such as GPS, GSM and possibly Wi-Fi, it should not be of concern for sound positioning as it cannot be used to determine position outside the areas equipped with the infrastructure.
2. Ultrasound trilateration is sufficient on its own and will not benefit much from merging with other positioning methods. Among GPS solutions only pseudolites work indoors, but they are currently not compatible with mobile phones. GSM provides no benefit, being insufficiently accurate and Bluetooth performs rather poorly with moving targets. Some simple form of Wi-Fi positioning may be used to track the user between locations for extra reliability. Considering a Wi-Fi connection will be needed anyway to send requests and content, this is not a major issue. Finally computer vision is a very promising solution on its own, but there is little benefit from combining it with sound trilateration. While computer vision can be very accurate, it will consume a lot of computational resources; require a lot of development and tweaking while at the same time being dependant on how the user physically operates the phone.
3. The ability to use ultrasound, which is inaudible to human ears, is an important attribute of a system that uses sound waves. If a sound signal used for trilateration was within the hearing range, it would appear sharp, loud, and overall unpleasant to human hearing. This is because a signal needs to be as distinct as possible in order to cover long distances, resist reverberation and clearly identify time-of-arrival. The concept is very similar to how fiducial markers in computer vision must be very vivid to allow accurate readings - unless the system uses infrared, which is invisible to human eyes.
4. Sound presents an effective way of using trilateration with conventional mobile phone hardware. Under the same temperature conditions, sound travels through air at a

constant and relatively slow speed. It is therefore possible to accurately deduce distance from time-of-arrival even at an average sample rate. In contrast, electromagnetic waves travel at the speed of light, so Wi-Fi, Bluetooth and GSM trilateration has to rely on signal strength, which is a much less reliable parameter.

5. Ultrasound positioning is compatible with many mobile interfaces. Because ultrasound positioning will work regardless of how the user holds the device, it is not restricted to a few applications, such as is the case with computer vision. At the same time, high accuracy of positioning means intelligent applications such as directional querying can be implemented. Finally ultrasound should not disrupt audio interfaces.

Our proposed approach is to generate a simple sine tone ultrasound signal using inbuilt mobile phone speakers. The signal is then received by up to four matched DPA microphones, each located in one corner of the test laboratory, and processed using a Pro Tools HD system. Live audio streams from the four microphones are then analyzed in real time by DSP filters tuned to specific ultrasound frequencies. The arrival time at each microphone is then used to calculate the position of the signal source using trilateration. The derived position can then be combined with accelerometer (pitch and roll) and magnetometer (yaw) readings (which are now standard on many smartphones) in order to obtain the position and orientation of the device. This combination of position and orientation can then be used for directional querying of specific points of interest (POI) within the environment, thus reducing the effect of display clutter or "information overload" on these small format devices. A Wi-Fi connection can be used to inform the server of the client's intention to send the tone, the tone's timestamp, the client's identity, plus any other information used by or for this application.

#### IV. EXPERIMENTAL DESIGN

In order to test the limitations of generating an ultrasound signal on mobile devices, experiments were carried out on a representative sample of four commercial off-the-shelf (COTS) smartphones: HTC G1, HTC Hero, Apple iPhone 3GS and Nokia 6210 Navigator.

First of all it was necessary to test the microphone which would be used to detect the signals. There are very few microphones that officially support frequencies up to 22 KHz. A majority of professional microphones officially cover 20 Hz to 20 KHz, with cheaper models sometimes stopping at 17 KHz. This is only a precaution as microphones are known to capture frequencies above the upper limit given in their specifications. So, since with microphones the specifications cannot be relied on, it is necessary to confirm that the chosen microphone can detect signals in the entire range, before each of the mobile phones can be tested.

In order to eliminate any incidental sounds, the experiments were done in a soundproof recording booth and with assistance from ProTools software. A Neumann U87 Ai microphone was successfully tested by playing one of the sound files, described later in this section, through Beyerdynamic DT150 earphones at high volume. The

specifications for these earphones state they can produce frequencies up to 30 KHz.

Initially one 44.1 KHz "WAV" sound file was generated using WaveLab software. This file starts with 10 seconds of silence in order to allow enough time to place the phone in front of the microphone, close the recording booth door and start recording. These ten seconds are followed by 11 one second long signals ranging from 17 to 22 KHz with a half KHz step. There is a gap of one second between each signal. A spectrogram of this file can be seen on Figure 3.

During the early stages of the experiment it was observed that mobile phones can generate a lot of noise in the lower frequencies when playing some or all of the given signals at maximum volume. This effect fades or disappears differently on different devices when volume is decreased. To counter this effect, the testing procedure was modified. First of all, four more modifications of the sound file were generated where volume is decreased by 20, 40, 60 and 80 percent. Secondly, each of the five files were played at maximum volume on the device as well as one and two steps lower from maximum. This resulted in 15 separate recordings per device or 60 altogether. A spectrogram was generated for each of the 60 recordings using Praat software for further analysis.

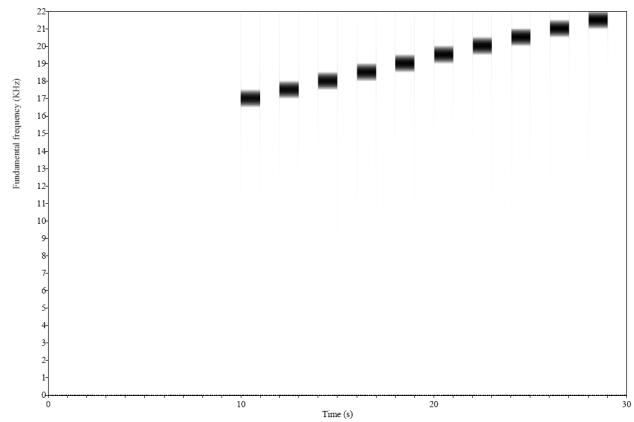


Figure 3. A spectrogram of the file played back by the smartphones. X axis depict time and Y axis depict Frequency. Chromatic value shows energy.

#### V. DISCUSSION

Based on the spectrograms generated during the experiment the following observations were made:

1. All tested devices are able to generate all of the given frequencies under the condition that the volume is not too high. In other words there was always energy in the part of the histogram corresponding to the signal. Also for every device it is possible to find a volume setting at which the spectrogram looks almost the same as the spectrogram of the original file. For example with G1 the settings will be file volume 80%, device volume maximum - 2 .(Figure 4)
2. If the volume is set too high, mobile phones will generate a lot of noise in a wide range of frequencies in the audible range when trying to generate one of the inaudible signals. For the iPhone, this happens only with 21.5 and 22 KHz, but for Hero and Navigator this happens at all tested frequencies. (See Figures 5 and 6.) Only HTC G1

appeared to be almost completely immune to this problem. As the volume is decreased, this problem fades, and at some point disappears. For example with HTC Hero this happens at around 80% file volume at maximum device volume. With iPhone noise at 21.5 and 22 KHz disappears completely around 20% file volume and device volume maximum - 2.

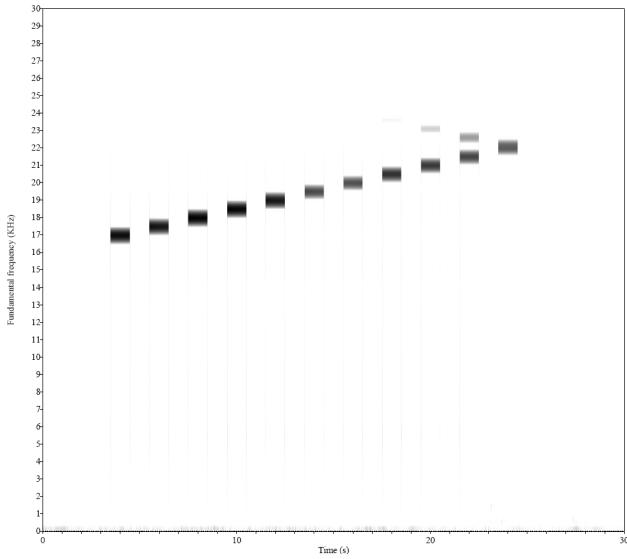


Figure 4. Spectrogram for HTC G1 at file volume 80, device volume maximum - 2

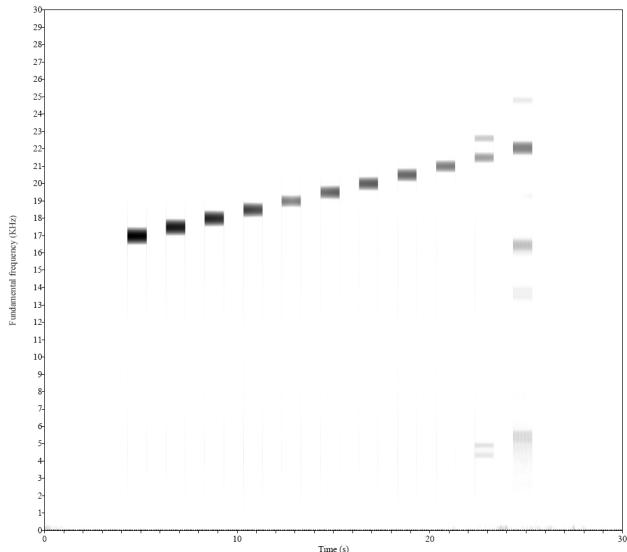


Figure 5. Spectrogram for iPhone at file volume 60%, device volume maximum

3. Volume settings of the device have a major impact on the appearance of noise. This was particularly evident with Nokia Navigator, where it was impossible to avoid noise even with 20% file volume. Noise almost completely disappeared when the device was set to maximum - 2 even with 100% file volume. With other devices it was only observed that noise can be almost completely eliminated by setting the device volume only one or two steps lower

than maximum. Reducing volume in the file seemed to have less impact. (See Figure 7 and 8 for comparison)

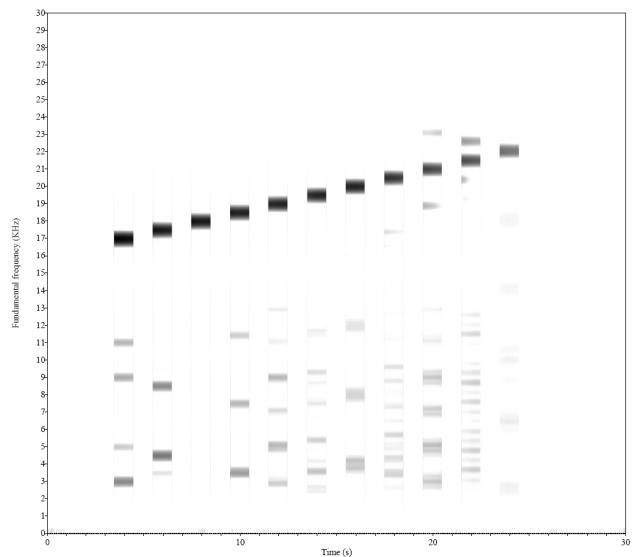


Figure 6. Spectrogram for HTC Hero at file volume 100%, device volume maximum

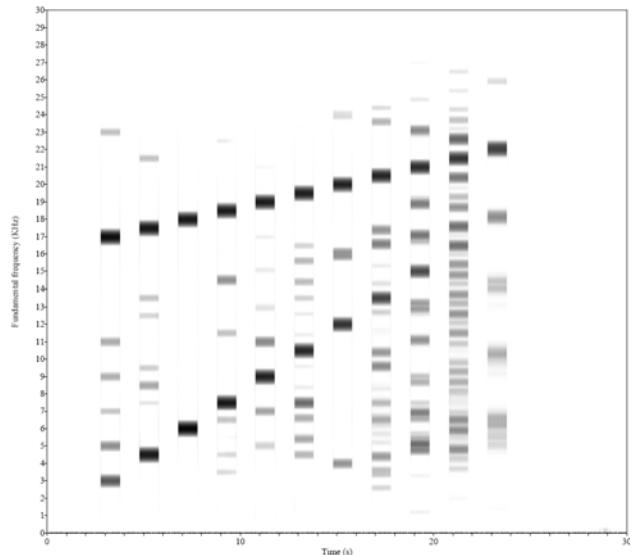


Figure 7. Spectrogram for Nokia Navigator at file volume 20%, device volume maximum. There is a lot of noise despite a very low volume playback of the signal in the file.

4. In a majority of recordings there can be observed a particular pattern of artefacts which are a few KHz higher than the real signals. Sometimes they are almost as powerful as the real signal, but very often are hardly visible. A very vivid example can be seen on Figure 8, but for other phones the effect is close to Figure 4. This is probably caused by either resonance in speaker diaphragm or operational errors in Digital Signal Processing (DSP) hardware. This trend may impact scalability of the positioning system. For example as can be seen on Figure 4, the system wouldn't be able to tell whether the original signal was 21.5 or 22.5 KHz. If two different devices used

these different frequencies to uniquely identify themselves, the system would fail to tell whether the two signals are an original and a phantom or two simultaneous signals from the two devices.

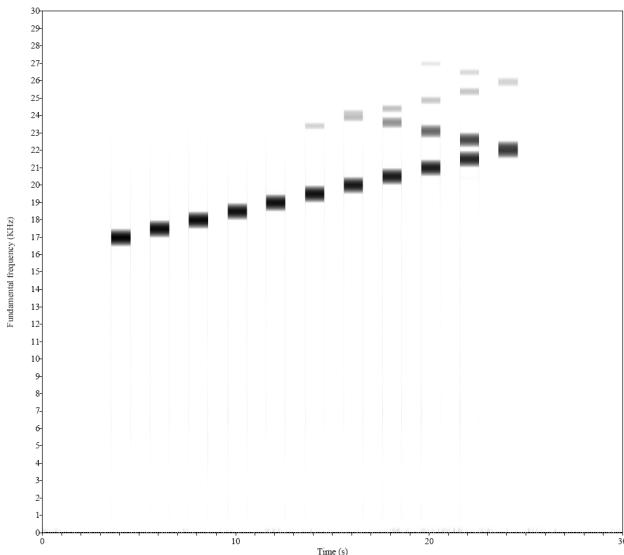


Figure 8. Spectrogram for Nokia Navigator at file volume 100%, device volume maximum - 2. Audible noise abruptly disappears at maximum - 2 settings even though file volume is high.

## VI. CONCLUSIONS

We presented some practical limitations of ultrasound generation on mobile phones. With the exception of very high volume settings, all tested mobile phones performed generation of 17-22 KHz signals very well. Some devices performed better than others. HTC G1 didn't generate almost any audible noise even at the highest settings. iPhone showed even less noise at the highest settings with the exception of 21.5 and 22 KHz signals. The other two phones generated a lot of noise at the highest volume settings. The problem with audible noise being generated along with ultrasound was easily avoided by reducing the volume settings on the device. Making the original signal quieter seemed to have less effect or even no effect at all on the Nokia 6210 Navigator. On most devices 20-22 KHz signals were accompanied by noise in the upper frequencies as on Figure 8. Reducing signal volume didn't have almost any effect on them. Although this noise is unavoidable it will not have any impact on usability, being inaudible. But it should be taken into consideration when scaling up the system to accommodate more devices. From our observations we can conclude that the cause of the noise in the upper frequencies is different from the cause of noise in lower frequencies.

None of the tested COTS devices met any overwhelming obstacles generating inaudible sound frequencies. Combined with what we learned from the literature, such as the findings of Peng et al., this shows mobile phone positioning using ultrasound trilateration as a promising direction for indoor LBS applications research.

## VII. FUTURE WORK

The following questions have to be answered next:

1. *What is the maximum distance at which an ultrasound signal emitted by a mobile phone can be reliably detected with a microphone?* Sound signals tend to fade with distance and even more so high-frequency signals. At the same time, if a signal is very loud it may get distorted by the microphone as well as being potentially audible to some people. Therefore an optimal volume must be found and the maximum distance at which the system can reliably distinguish it from background noise will be the maximum detection range.
2. *Can background noise cause false positives and how can this be countered?* There is a possibility that some electric device (e.g. router, network switch, air conditioner, power adapter etc.) in the room produces sound of the same frequency as the signal used by the positioning system and therefore regularly or irregularly causes the system to "detect" a false signal.
3. *What kind of ultrasound signal suffers the least from multipath and reverberation under normal room conditions?* There are a number of signal properties to experiment with such as volume, frequency, length and shape (e.g. linear increase/decrease of amplitude).
4. *What accuracies can ultrasound trilateration offer?* First of all it must be found with what accuracy the distance between one speaker and one microphone can be detected. Secondly with what accuracy a mobile device can be located in a 2D plane using an array of microphones. And finally with what accuracy a mobile phone can be located in three dimensions.
5. *How can a digital compass be configured to give accurate readings indoors?* While this question is not directly linked to positioning, it needs to be answered in order to test how well the proposed method performs for directional querying. Magnetometers are easily distorted by local magnetic fields, which are abundant indoors. They are however expected to exhibit the same deviations in the same locations, so it may be possible to improve accuracy through the process of "weighting", considering accurate position is available.
6. *Can the combination of ultrasound positioning and readings from accelerometers and a digital compass be combined to allow for directional querying?* This will help evaluate how well the proposed method performs a useful indoor LBS task.

## ACKNOWLEDGMENT

The authors wish to thank the Higher Education Authority (HEA) in Ireland and specifically their Technological Sector Research Strand III: Core Research Strengths Enhancement Programme for funding the work carried out at the Dublin Institute of Technology on the Lok8 project.

## REFERENCES

- [1] "Ekahau RTLS Overview.", Ekahau.com, <http://www.ekahau.com/products/real-time-location-system/overview.html>, 21 May, 2009.
- [2] M. Meijers, Zlatanova, S., Pfeifer, N., "3D Geo-Information Indoors: Structuring for Evacuation," in *First International Workshop on Next Generation 3D City Models*, 2005, pp. 11-16.
- [3] "Point Inside.", pointinside.com, <http://www.pointinside.com>, 8 May, 2010

- [4] B. Siciliano, Khatib, O., *Springer Handbook of Robotics*: Springer, 2008.
- [5] B. Williams, Klein, G., Reid, I. , "Real-Time SLAM Relocalisation," in *Computer Vision*, 2007.
- [6] D. Wagner, Schmalstieg, D., "First Steps Towards Handheld Augmented Reality," in *7th IEEE International Symposium on Wearable Computers*, 2003, pp. 127-136.
- [7] V. Otsason, Varshavsky, A., LaMarca, A., De Lara, E., "Accurate GSM Indoor Localization," in *Pervasive and Mobile Computing*, 2007.
- [8] B. Ferris, Hähnel, D., Fox, D, "Gaussian Processes for Signal Strength-Based Location Estimation," in *Robotics Science and Systems* 2006.
- [9] "Wi-Fi Alliance announces groundbreaking specification to support direct Wi-Fi connections between devices.", wi-fi.org, [http://www.wi-fi.org/news\\_articles.php?f=media\\_news&news\\_id=909](http://www.wi-fi.org/news_articles.php?f=media_news&news_id=909), 10 November, 2009.
- [10] S. Zhou, Pollard, J., "Position measurement using Bluetooth," *Consumer Electronics, IEEE Transactions on*, vol. 52, pp. 555-558, May 2006.
- [11] K. Kolodziej, Hjelm, J., *Local positioning systems : LBS applications and services*, 2006.
- [12] M. Addlesee, Curwen, R., Hodges, S., Newman, J., Steggles, P., Ward, A., Hopper, A., "Implementing a Sentient Computing System," *IEEE Computer*, vol. 34, pp. 50-56, August 2001.
- [13] J. Hallberg, Nilsson, M., Synnes, K., "Positioning with Bluetooth," in *ICT*, 2003.
- [14] G. Borriello, Liu, Alan., Offer, T., Palistrant, C., Sharp, R., "WALRUS: Wireless Acoustic Location with Room-Level Resolution using Ultrasound," in *Mobisys*, 2005, pp. 191-203.
- [15] C. Peng, Shen, G., Zhang, Y., Li, Y., Tan, K., "BeepBeep: A High Accuracy Acoustic Ranging System using COTS Mobile Devices," in *SenSys*, 2007.
- [16] "Navizon Technical Paper", navizon.com, [http://www.navizon.com/Navizon\\_wifi\\_gps\\_and\\_cell\\_tower\\_positioning.pdf](http://www.navizon.com/Navizon_wifi_gps_and_cell_tower_positioning.pdf), 11 January, 2009.
- [17] C. Hsiao, Huang, P., "Two Practical Considerations of Beacon Deployment for Ultrasound-Based Indoor Localization Systems," in *Sensor Networks, Ubiquitous and Trustworthy Computing*, 2008, pp. 306-311.
- [18] "Skyhook Wireless: How it works", skyhookwireless.com, <http://www.skyhookwireless.com/howitworks>, 12 May, 2009.

# **Asynchronous Ultrasonic Trilateration for Indoor Positioning of Mobile Phones**

Viacheslav Filonenko, Charlie Cullen and James D. Carswell

Web and Wireless Geographical Information Systems

(W2GIS 2012)

12 – 13 April 2012,

Naples, Italy

Pages 33-46

Volume 7236 of Lecture Notes in Computer Science

Springer-Verlag, 2012

# Asynchronous Ultrasonic Trilateration for Indoor Positioning of Mobile Phones

Viacheslav Filonenko, Charlie Cullen, and James D. Carswell

Digital Media Centre, Dublin Institute of Technology, Ireland

{viacheslav.filonenko,charlie.cullen,jcarswell}@dit.ie

**Abstract.** In this paper we discuss how the innate ability of mobile phone speakers to produce ultrasound can be used for accurate indoor positioning. The frequencies in question are in a range between 20 and 22 KHz, which is high enough to be inaudible by humans but still low enough to be generated by today's mobile phone sound hardware. Our tests indicate that it is possible to generate the given range of frequencies without significant distortions, provided the signal volume is not turned excessively high. In this paper we present and evaluate the accuracy of our asynchronous trilateration method (Lok8) for mobile positioning without requiring knowledge of the time the ultrasonic signal was sent. This approach shows that only the differences in time of arrival to multiple microphones (control points) placed throughout the indoor environment is sufficient. Consequently, any timing issues with client and server synchronization are avoided.

**Keywords:** Indoor Mobile Positioning, Ultrasonic Trilateration, LBS.

## 1 Introduction

The role of mobile phones in society has changed dramatically in the past few years as for many people their SmartPhone is an omnipresent gateway to information. The mobile nature of the device is of key importance. Users have come to expect constant access to the phone's information facilities in many different circumstances and environments that take into account location and personal preference when providing useful and timely decision support services.

Currently outdoor Location Based Services (LBS) have the advantage of relatively reliable positioning via GPS (also Wi-Fi and GSM) and a defined business model for the delivery of content to the user. This has led to applications of outdoor LBS greatly expanding in recent years, leaving indoor locationing technologies and services on mobile devices to yet fully mature. The current state-of-the-art of merging accurate (i.e. sub-metre) indoor positioning and context-sensitive services for indoor LBS therefore is still an open problem.

The following factors make indoor positioning challenging:

1. Generally indoor environments require higher accuracy to be useful for practical LBS purposes. This is because when indoors we are dealing with objects and distances on a smaller scale. While accuracy of +/- 10 meters may be good enough

to direct someone to a cafe or a bus stop, indoors it could mean we are unsure in which room the user currently is located.

2. Locationing services that rely on satellite signals such as GPS for positioning do not work indoors at all because these signals require a direct line-of-sight to the receiver.
3. When used indoors, electromagnetic fields and sound signals can suffer from fading and multipath propagation when they encounter walls, windows, and other structures. This requires implementations of a robust solution that can effectively overcome the positioning difficulties typically found in cluttered, complex indoor environments.

Under these circumstances it is understandable that very specialized hardware may be required unless we are willing to sacrifice accuracy. However, given the role of commercial off-the-shelf (COTS) SmartPhones in today's society, they have by default become the platform of choice for implementations of indoor positioning and therefore the standard hw platform we have developed our Lok8 (locate) indoor positioning solution to work on.

While many mobile positioning approaches are erroneously described in the press as *triangulation*, where angles between mobile devices to various receivers (control points) would be required, what is in fact being described is *trilateration*, where distances to known control points or beacons are instead used in the positioning calculation. Significantly, what is often common among these solutions is some sort of timing synchronisation requirement between transmitter and receiver to provide a full measure of distance as inputs to the trilateration process.

The main contribution of our Lok8 approach is that we remove this often delicate synchronisation problem between transmitter and receiver by instead requiring that receivers (i.e. 3 or more microphones) be connected to a central server that starts a timer once an ultrasonic signal is detected by any of the mics. By making the time the original signal was sent irrelevant, only the differences in time between when the signal reaches each of the remaining microphones is needed in our solution. The result is a more robust mobile trilateration method.

As comprehensive explanation of mobile trilateration procedure is all too rare in indoor LBS literature, another worthwhile contribution of this paper is to describe in detail our subtle but significant modification to standard least squares trilateration in Section 3. Where standard trilateration assumes that distances from an unknown position to all control points are known a-priori, instead we only know the differences between these distances - not the distances themselves. So while our asynchronous trilateration derivation is similar to the standard case, the initial conditions are different and therefore the standard trilateration solution requires modification. Before this we first discuss some background work in Section 2, and follow this with a summary of our principal contributions to the field of indoor mobile positioning and plans for future work in Section 4.

## 2 Indoor Positioning Background

There are many different methods and reported accuracies for locating a mobile device indoors (see Table 1). Methods that use propagation of Radio Frequency (RF) signals are prevalent in this field, with the exception of computer vision, where simultaneous localization and mapping (SLAM) appears to be the most promising but considered by many an operational technology still in its infancy [1]. Computer vision techniques, while potentially very accurate, is characterized by high computational load, complicated procedures of recovery from tracking failures and susceptibility to camera shake and motion blur. These problems are addressed in the studies done by Williams et al. [2] and Wagner et al. [3]. Another difficulty associated with computer vision is that the user is supposed to be looking through the display screen when using the device.

**Table 1.** Comparison of Indoor Positioning Implementations

	Best Accuracy	Underlying Technology	Available on SmartPhones
<b>Wide Signal Strength Fingerprinting</b>	2.48m	GSM	no
<b>Skyhook(GSM)</b>	200m	GSM	yes
<b>Navizon(GSM)</b>	50m	GSM	yes
<b>Skyhook(Wi-Fi)</b>	10m	Wi-Fi	yes
<b>Navizon(Wi-Fi)</b>	20m	Wi-Fi	yes
<b>RADAR</b>	2m	WaveLan	no
<b>GP for Signal Strength-Based Location Estimation</b>	2m	Wi-Fi	yes
<b>Ekahau</b>	1m	Wi-Fi	no
<b>The Bat</b>	3cm	Ultrasound	no
<b>The Cricket</b>	3cm	Ultrasound	no
<b>Lok8</b>	<b>Sub-metre</b>	<b>Ultrasound</b>	yes

RF-based transceivers such as GSM, Wi-Fi and Bluetooth can be found on every modern SmartPhone. Five meter accuracy, one of the best results for indoor GSM positioning, was displayed by Otsason et al. with the help of wide signal-strength fingerprinting [4]. Unfortunately wide signal-strength fingerprinting is impossible on many modern phones due to OS restrictions. Other GSM positioning methods are generally impractical for indoor use due to poor accuracy. Wi-Fi positioning is on average better than twice as accurate as GSM. A method proposed by Ferris et al. where Gaussian processes are used to mathematically predict signal strength in areas outside the exact spots where fingerprints were taken seems to promising [5]. The best accuracy among commercial solutions using this approach was shown by Ekahau: 1-3 meters [6].

Bluetooth has the shortest range among the three wireless technologies but there are two major problems that make Bluetooth positioning particularly difficult. First of all it is designed to adjust signal strength when signals become too strong or too weak making any subsequent distance measurements based on signal strength unreliable. Disabling this feedback loop is discussed by Zhou et al. [7]. Another problem is that it takes a lot of time for a new device to be fully discovered. Very often it means that the user has already left the area [8]. This makes Bluetooth trilateration impractical; however coarser room-level positioning can be done relatively quickly as device pairing is not required.

Notably, it is not reported possible to achieve accuracy below one meter [9] using RF-based technologies present in mobile phones [4, 5, 10]. However, sound travels at significantly slower speeds than radio waves and can therefore be easily localised to a few centimetres due to this much longer time of flight. Other useful features of sound show that it is possible to emit a 21 KHz (just above the human hearing range) signal from a mobile phone speaker and successfully receive it with a conventional microphone [11]. In a separate study, Peng et al. [12] showed that it is possible to utilize sound in order to measure the distance between two mobile phones using synchronized time-of-arrival techniques.

In previous work [16], we tested the useable range of SmartPhone ultrasound to find that these signals can indeed be successfully detected up to distances of 20m or more (Figure 1). In this experiment, two values below 10 dB were registered but this is still well above the 21.5 KHz component of background noise, which is around 1 dB. However there is no guarantee that the maximum value belongs to a signal that arrived by direct path and not via a longer deflected path. In any case, it can be seen from the shape of the graph that even with speaker and microphone pointing directly at each other, signal strength can't be relied on alone to accurately measure distance.



**Fig. 1.** Relationship between signal strength and distance for conditions where SmartPhone speaker and microphone point at each other

Therefore, in our Lok8 trilateration method we endeavour to make use of the useful characteristics of inaudible mobile ultrasound by exploiting the *differences* in signal time-of-arrival at a static microphone array for accurate mobile positioning. An accuracy comparison of our method compared to other reported indoor positioning methods, together with their availability for implementation on today's SmartPhones, is also given in Table 1.

### 3 Time Difference of Arrival (TDOA) Trilateration

Sound is a mechanical wave which travels at speeds much slower than the speed of light. In dry air at a temperature of 25°C the speed of sound is only 346 m/s. At such propagation speeds, one sample of a standard 44.1 KHz stream (44100 cycles/second) accounts for 0.8cm of distance [4, 13]. In other words a signal will travel only 0.8 centimeters in the duration of the smallest time grain. Technically it is possible to work with sound even at 384 KHz, which can give much finer accuracy.

As discussed previously, by using trilateration it is possible to calculate one's position based on the distance to several other (control) points with known positions [14, 15]. To find one's position in 2 dimensions the number of required known points is 3; for position in 3 dimensions the number of known points is 4. Given that the speed of sound propagation is constant under the same temperature and humidity conditions, the time it takes a signal to travel between the phone to each known microphone control point can be directly converted into distance between the phone and microphones. This is the TOA (Time of Arrival) approach. In general, the main problem with this approach is that both the time the signal was sent and the time it was received are required in order to get the time of flight.

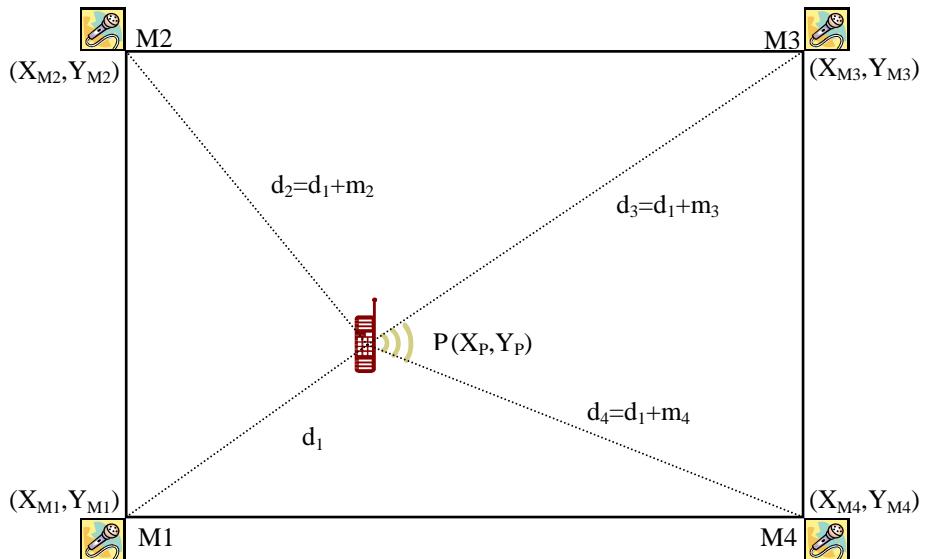
In our scenario of quickly and accurately locating a mobile phone indoors, TOA requires that times from two separate systems with two separate clocks will have to be synchronised - a major source of error. As such it is desirable to compare only the time of arrival at each of the microphones and ignore completely the time the signal was originally sent from the phone, making Lok8 a TDOA (Time Difference of Arrival) approach. The problem is illustrated in Figure 2 and the detailed solution follows.

#### Problem

- Mobile phone (P) has unknown position ( $X_p, Y_p$ ).
- 4 microphones (M1, M2, M3, M4) have known positions ( $X_{M1}, Y_{M1}$ ), ( $X_{M2}, Y_{M2}$ ), ( $X_{M3}, Y_{M3}$ ), ( $X_{M4}, Y_{M4}$ )
- 4 distances ( $d_1, d_2, d_3, d_4$ ) from P to M1, M2, M3, M4 are unknown but the differences between them ( $m_2, m_3, m_4$ ) are measured ultrasonically; these are the *observations*.
- Find coordinates of  $P=(X_p, Y_p)$  by solving a system of equations (mathematical model) that relates the  $m = 3$  *observations* ( $m_2, m_3, m_4$ ) to the  $n = 2$  *unknown parameters* ( $X_p, Y_p$ ).

## Solution

Although the coordinates of P could be found using readings from only 3 microphones (2 observations), 4 or more readings can be effectively used in the method of *Least Squares* to determine the *Most Probable Value* (MPV) for the coordinates of P, plus a *Standard Deviation* for the MPV.



**Fig. 2.** Time Difference of Arrival. Control points  $M_1$ ,  $M_2$ ,  $M_3$  and  $M_4$  are known microphone positions. Point  $P$  is the unknown mobile phone's position, coordinates of which we are trying to find. Lines  $d_1$ ,  $d_2$ ,  $d_3$  and  $d_4$  are unknown distances between the phone and each microphone. However, what are known are the differences between the three measurements:  $m_2$ ,  $m_3$  and  $m_4$ .

## Least Squares Method for TDOA Trilateration

From Pythagoras we derive the following *mathematical model* to describe the ultrasonic relationships between phone P and microphones  $M_1$ ,  $M_2$ ,  $M_3$ ,  $M_4$ :

$$\begin{aligned} d_1^2 &= (X_P - X_{M1})^2 + (Y_P - Y_{M1})^2 \quad \text{or} \quad d_1 = \sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2} \\ d_2^2 &= (X_P - X_{M2})^2 + (Y_P - Y_{M2})^2 \quad \text{or} \quad d_2 = \sqrt{(X_P - X_{M2})^2 + (Y_P - Y_{M2})^2} \\ d_3^2 &= (X_P - X_{M3})^2 + (Y_P - Y_{M3})^2 \quad \text{or} \quad d_3 = \sqrt{(X_P - X_{M3})^2 + (Y_P - Y_{M3})^2} \\ d_4^2 &= (X_P - X_{M4})^2 + (Y_P - Y_{M4})^2 \quad \text{or} \quad d_4 = \sqrt{(X_P - X_{M4})^2 + (Y_P - Y_{M4})^2} \end{aligned}$$

However, we can re-write  $d_2$ ,  $d_3$ ,  $d_4$  in terms of  $d_1$ :

$$\begin{aligned}d_2 &= d_1 + m_2 \\d_3 &= d_1 + m_3 \\d_4 &= d_1 + m_4\end{aligned}$$

And then substitute above  $d_1$  expressions back into the mathematical model:

$$\begin{aligned}d_1 + m_2 &= \sqrt{(X_P - X_{M2})^2 + (Y_P - Y_{M2})^2} \quad \text{or } m_2 = \sqrt{(X_P - X_{M2})^2 + (Y_P - Y_{M2})^2} - d_1 \\d_1 + m_3 &= \sqrt{(X_P - X_{M3})^2 + (Y_P - Y_{M3})^2} \quad \text{or } m_3 = \sqrt{(X_P - X_{M3})^2 + (Y_P - Y_{M3})^2} - d_1 \\d_1 + m_4 &= \sqrt{(X_P - X_{M4})^2 + (Y_P - Y_{M4})^2} \quad \text{or } m_4 = \sqrt{(X_P - X_{M4})^2 + (Y_P - Y_{M4})^2} - d_1\end{aligned}$$

Then replace  $d_1$  in  $m_2$ ,  $m_3$ ,  $m_4$  equations above with equivalent  $d_1$  expression from mathematical model to give:

$$\begin{aligned}m_2 &= \sqrt{(X_P - X_{M2})^2 + (Y_P - Y_{M2})^2} - \sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2} \\m_3 &= \sqrt{(X_P - X_{M3})^2 + (Y_P - Y_{M3})^2} - \sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2} \\m_4 &= \sqrt{(X_P - X_{M4})^2 + (Y_P - Y_{M4})^2} - \sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2}\end{aligned}$$

Re-write above three mathematical model equations as *observation equations* by adding a residual  $v_m$  to each measurement:

$$\begin{aligned}F: \quad m_2 + v_{m2} &= \sqrt{(X_P - X_{M2})^2 + (Y_P - Y_{M2})^2} - \sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2} \\G: \quad m_3 + v_{m3} &= \sqrt{(X_P - X_{M3})^2 + (Y_P - Y_{M3})^2} - \sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2} \\H: \quad m_4 + v_{m4} &= \sqrt{(X_P - X_{M4})^2 + (Y_P - Y_{M4})^2} - \sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2}\end{aligned}$$

Because number of measurements ( $m = 3$ ) is greater than number of unknowns ( $n = 2$ ), use Least Squares to determine the MPV of the unknowns ( $X_P, Y_P$ ). Since the observation equations are non-linear in the unknowns ( $X_P, Y_P$ ), a first-order *Taylor Series* is needed to approximate a set of linear observation equations before taking partial derivatives.

Considering function F above (describing ultrasonic relationship between M2 and P):

$$F: \quad m_2 + v_{m2} = \sqrt{(X_P - X_{M2})^2 + (Y_P - Y_{M2})^2} - \sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2}$$

This non-linear function can be written as:

$$F(X_P, Y_P) = m_2 + v_{m2}$$

Where

$$F(X_P, Y_P) = \sqrt{(X_P - X_{M2})^2 + (Y_P - Y_{M2})^2} - \sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2}$$

The above function is *linearized* using a first-order Taylor Series approximation:

$$F(X_P, Y_P) = F(X_{Po}, Y_{Po}) + \left( \frac{\partial F}{\partial X_P} \right)_o dX_P + \left( \frac{\partial F}{\partial Y_P} \right)_o dY_P$$

Where

- $X_{Po}$  and  $Y_{Po}$  are initial estimates of SmartPhone position in the environment calculated by taking average of all known microphone positions.
- $F(X_{Po}, Y_{Po})$  is the non-linear function evaluated with these estimates.
- $dX_P$  and  $dY_P$  are corrections to the initial estimates such that  $X_p = X_{p_o} + dX_p$  and  $Y_p = Y_{p_o} + dY_p$ .

The partial derivatives  $\left( \frac{\partial F}{\partial X_P} \right)$  and  $\left( \frac{\partial F}{\partial Y_P} \right)$  are found by first re-writing function F:

$$F: \quad F(X_P, Y_P) = \left( (X_P - X_{M2})^2 + (Y_P - Y_{M2})^2 \right)^{\frac{1}{2}} - \left( (X_P - X_{M1})^2 + (Y_P - Y_{M1})^2 \right)^{\frac{1}{2}}$$

and then take partial derivative with respect to  $X_P$ :

$$\begin{aligned} \frac{\partial F}{\partial X_P} &= \frac{1}{2} \left( (X_P - X_{M2})^2 + (Y_P - Y_{M2})^2 \right)^{-\frac{1}{2}} \bullet 2(X_P - X_{M2}) \\ &\quad - \frac{1}{2} \left( (X_P - X_{M1})^2 + (Y_P - Y_{M1})^2 \right)^{-\frac{1}{2}} \bullet 2(X_P - X_{M1}) \\ &= \frac{(X_P - X_{M2})}{\sqrt{(X_P - X_{M2})^2 + (Y_P - Y_{M2})^2}} - \frac{(X_P - X_{M1})}{\sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2}} \\ &= \frac{(X_P - X_{M2})}{d_1 + m_2} - \frac{(X_P - X_{M1})}{d_1} \end{aligned}$$

and then with respect to  $Y_P$ :

$$\begin{aligned}\frac{\partial F}{\partial Y_P} &= \frac{1}{2} \left( (X_P - X_{M2})^2 + (Y_P - Y_{M2})^2 \right)^{\frac{1}{2}} \bullet 2(Y_P - Y_{M2}) \\ &\quad - \frac{1}{2} \left( (X_P - X_{M1})^2 + (Y_P - Y_{M1})^2 \right)^{\frac{1}{2}} \bullet 2(Y_P - Y_{M1}) \\ &= \frac{(Y_P - Y_{M2})}{\sqrt{(X_P - X_{M2})^2 + (Y_P - Y_{M2})^2}} - \frac{(Y_P - Y_{M1})}{\sqrt{(X_P - X_{M1})^2 + (Y_P - Y_{M1})^2}} \\ &= \frac{(Y_P - Y_{M2})}{d_1 + m_2} - \frac{(Y_P - Y_{M1})}{d_1}\end{aligned}$$

Where  $d_1$  is always (re)evaluated using Pythagoras at current estimates for  $(X_P, Y_P)$ .

Therefore:

$$\begin{aligned}F(X_P, Y_P) &= F(X_{Po}, Y_{Po}) + \left( \frac{(X_P - X_{M2})}{d_1 + m_2} - \frac{(X_P - X_{M1})}{d_1} \right)_o dX_P \\ &\quad + \left( \frac{(Y_P - Y_{M2})}{d_1 + m_2} - \frac{(Y_P - Y_{M1})}{d_1} \right)_o dY_P\end{aligned}$$

So the linearized observation equation for  $m_2$ , describing the ultrasonic relationship between microphone M2 and phone P becomes:

$$\begin{aligned}&\left( \frac{(X_P - X_{M2})}{d_1 + m_2} - \frac{(X_P - X_{M1})}{d_1} \right)_o dX_P + \left( \frac{(Y_P - Y_{M2})}{d_1 + m_2} - \frac{(Y_P - Y_{M1})}{d_1} \right)_o dY_P \\ &= (m_2 - m_{2o}) + v_{m2}\end{aligned}$$

Likewise for function G (between M3 and P):

$$\begin{aligned}&\left( \frac{(X_P - X_{M3})}{d_1 + m_3} - \frac{(X_P - X_{M1})}{d_1} \right)_o dX_P + \left( \frac{(Y_P - Y_{M3})}{d_1 + m_3} - \frac{(Y_P - Y_{M1})}{d_1} \right)_o dY_P \\ &= (m_3 - m_{3o}) + v_{m3}\end{aligned}$$

and function H (between M4 and P):

$$\begin{aligned}&\left( \frac{(X_P - X_{M4})}{d_1 + m_4} - \frac{(X_P - X_{M1})}{d_1} \right)_o dX_P + \left( \frac{(Y_P - Y_{M4})}{d_1 + m_4} - \frac{(Y_P - Y_{M1})}{d_1} \right)_o dY_P \\ &= (m_4 - m_{4o}) + v_{m4}\end{aligned}$$

When using Matrix Methods for Least Squares, the observation equations are represented in matrix form as:

$${}_m A_{nn} X_1 = {}_m L_1 + {}_m V_1$$

Where in our case:

- $m = 3, n = 2$
- ${}_m A_n$  contains the coefficients of the unknowns  $(X_P, Y_P)$
- ${}_n X_1$  contains the corrections to be applied to the initial estimates for the unknowns  $(dX_P, dY_P)$
- ${}_m L_1$  contains the measurements  $(m_2, m_3, m_4)$
- ${}_m V_1$  contains the residuals (one for each measurement).

Solving for  $X$  gives the solution:

$$X = (A^T A)^{-1} A^T L \text{ where:}$$

$$A = \begin{bmatrix} \frac{(X_P - X_{M2})}{d_1 + m_2} - \frac{(X_P - X_{M1})}{d_1} & \frac{(Y_P - Y_{M2})}{d_1 + m_2} - \frac{(Y_P - Y_{M1})}{d_1} \\ \frac{(X_P - X_{M3})}{d_1 + m_3} - \frac{(X_P - X_{M1})}{d_1} & \frac{(Y_P - Y_{M3})}{d_1 + m_3} - \frac{(Y_P - Y_{M1})}{d_1} \\ \frac{(X_P - X_{M4})}{d_1 + m_4} - \frac{(X_P - X_{M1})}{d_1} & \frac{(Y_P - Y_{M4})}{d_1 + m_4} - \frac{(Y_P - Y_{M1})}{d_1} \end{bmatrix}$$

$$X = \begin{bmatrix} dX_P \\ dY_P \end{bmatrix} \quad L = \begin{bmatrix} m_2 - m_{2_0} \\ m_3 - m_{3_0} \\ m_4 - m_{4_0} \end{bmatrix} \quad V = \begin{bmatrix} v_{m2} \\ v_{m3} \\ v_{m4} \end{bmatrix}$$

Matrix  $X$  contains the corrections to be applied to the original estimates for  $(X_P, Y_P)$ . These new  $(X_P, Y_P)$  coordinates are then used to recalculate updated distances for  $(d_1, m_{2_0}, m_{3_0}, m_{4_0})$ . The process is repeated until coordinates of  $(X_P, Y_P)$  don't change significantly (e.g. in the 3<sup>rd</sup> decimal place for mm precision).

After a solution has been reached, the residuals  $V$  for each measurement and Standard Deviation of *unit weight*  $\sigma_o$  for the overall least squares adjustment can be calculated with:

$$V = AX - L \quad \text{and} \quad \sigma_o = \pm \sqrt{\frac{(V^T V)}{r}}$$

Where *degrees of freedom*  $r = m-n$  and the Standard Deviation of each adjusted unknown is then given by:

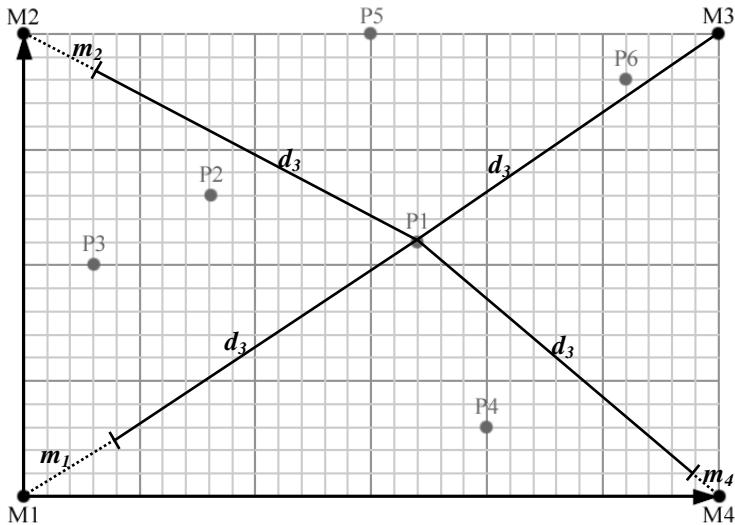
$$\sigma_{Xi} = \pm \sigma_o \sqrt{(Q_{XiXi})}$$

In our case  $\sigma_{X_1}$  is the Standard Deviation for  $X_P$ , and  $\sigma_{X_2}$  is the Standard Deviation for  $Y_P$ . These standard deviations imply that there is a 68% probability that the adjusted values for  $X_P$  and  $Y_P$  are within  $\pm\sigma$  of this amount.

$(A^T A)^{-1}$  is called the *variance-covariance* matrix or  $(Q_{XX})$  matrix and  $(Q_{x_i x_i})$  is the *variance* of unknown  $i$ , or the element in the  $i^{th}$  row and  $i^{th}$  column of the  $(A^T A)^{-1}$  matrix.

### Practical Example

To test the accuracy of our TDOA Trilateration method, we used it to calculate the position of several random SmartPhone locations and compare the results to their actual positions in Figure 3. We used four control points (microphones) arranged in the corners of a rectangular room to locate the phone's position at 6 different locations within the room.



**Fig. 3.** TDOA Trilateration experiment with four microphones and six different smartphone positions. Control points M1, M2, M3 and M4 are microphones. Points P1, P2, P3, P4, P5 and P6 are actual SmartPhone locations. Each square of the grid represents 1 unit in length.

Regarding input data for testing the Lok8 trilateration algorithm, the locations of M1(0,0), M2(0,20), M3(30,20), M4(30,0) were used and the initial distances between the mics and the various phone positions were measured manually. Although we could have used Pythagoras in Figure 3 to calculate exactly the measurements representing the ultrasonic distances between the microphones and various phone positions, we wanted to introduce some error in the measurements so chose instead to simply use a ruler to measure these distances on paper to one decimal point precision. After that we subtracted the shortest measured distance for any given phone position from each of the remaining three mic distances. The resulting 3 distance differences

were then used as “ultrasonic” input to the asynchronous trilateration procedure in addition to the known microphone locations.

For example, for phone position P1 the measured distance to M1 was 20.2, to M2 19.2, M3 15.8, and M4 17.0. The shortest distance is to M3, therefore it is subtracted from the other 3 distances to leave;  $m_1 = 4.4$ ,  $m_2 = 3.4$ ,  $m_4 = 1.2$ . These values simulate time measurements translated to distance for the ultrasonic signal to reach these 3 mics after first triggering the server clock at M3. The input data is summarised in Table 2 and the trilateration results for the phone’s position relative to the 4 microphones are compiled in Table 3. Notice that if we assumed metres for units in this example, the standard deviations for the phone positions are of sub-metre accuracy.

**Table 2.** Sample TDOA Trilateration input. Second and third columns contain coordinates of a microphone and fourth column contains differences between distance to mic and closest mic. In this example microphone M3 is closest to phone position P1 so its corresponding distance difference equals zero.

Mic	X	Y	Distance Difference ( $m_i$ )
M1	0	0	4.4
M2	0	20	3.4
M3	30	20	0
M4	30	0	1.2

**Table 3.** Comparison of TDOA output and expected results. Second column contains X and Y coordinates of a given phone position, third column contains coordinates of the phone as calculated by our TDOA trilateration procedure. Fourth and fifth columns contain the Standard Deviations ( $\sigma_x, \sigma_y$ ) for each trilaterated phone positon and number of iterations to get there.

Phone Point	Actual Location	TDOA Trilateration	Standard Deviation	Number of Iterations
P1	17 , 11	16.987, 10.986	0.0002, 0.0003	3
P2	8 , 13	7.978, 12.966	0.0158, 0.019	3
P3	3 , 10	2.96, 10.0	0 , 0	4
P4	20 , 3	20.002, 2.996	0.011, 0.0195	3
P5	15 , 20	15.0, 20.0	0 , 0	4
P6	26 , 18	25.999, 18.031	0.0144, 0.0214	4

## 4 Conclusions and Future Work

In this paper we demonstrated an asynchronous trilateration method that can be reliably used to accurately locate an ultrasonic signal source without knowing the time the signal was sent. This eliminates the need to synchronize clocks between signal source and receivers.

An advantage of using a Least Squares approach for trilateration is its ability to tolerate errors in measurements; with more measurements provided, less is the impact from a single erroneous measurement. Also, due to the iterative nature of this approach allowing for a large *pull-in range*, initial approximations for a phone's position in a room can be simply taken as the average of all microphone (control point) positions. While the algorithm can work with only three receivers (mics), at least four or more are recommended for scenarios where measurements are likely to be contaminated with signal noise caused by multipath propagation.

For future work we plan to implement our TDOA Trilateration method in a real-time indoor positioning system on COTS SmartPhones and interconnected mics. We will then evaluate how well Lok8 manages with unavoidable measurement errors due to background noise, obstructions, and uncertainty due to the presence of multiple ultrasonic source devices.

**Acknowledgments.** The authors wish to thank the Higher Education Authority (HEA) in Ireland and their Technological Sector Research Strand III for funding the ultrasound work in the Lok8 project. Preparation of this publication and the asynchronous trilateration work was funded by a Strategic Research Cluster Grant (07/SRC/I1168) by Science Foundation Ireland under the National Development Plan. The authors also gratefully acknowledge this support.

## References

1. Siciliano, B., Khatib, O.: Springer Handbook of Robotics. Springer, Heidelberg (2008)
2. Williams, B., Klein, G., Reid, I.: Real-Time SLAM Relocalisation. In: Computer Vision (2007)
3. Wagner, D., Schmalstieg, D.: First Steps Towards Handheld Augmented Reality. In: 7th IEEE International Symposium on Wearable Computers. IEEE Computer Society (2003)
4. Otsason, V., Varshavsky, A., LaMarca, A., De Lara, E.: Accurate GSM Indoor Localization. In: Pervasive and Mobile Computing (2007)
5. Ferris, B., Hähnel, D., Fox, D.: Gaussian Processes for Signal Strength-Based Location Estimation. In: Robotics Science and Systems (2006)
6. Ekahau RTLS Overview, <http://www.ekahau.com/products/real-time-location-system/overview.html> (cited May 21, 2009)
7. Zhou, S., Pollard, J.: Position measurement using Bluetooth. IEEE Transactions on Consumer Electronics 52(2), 555–558 (2006)
8. Kolodziej, K., Hjelm, J.: Local positioning systems: LBS applications and services (2006)
9. Addlesee, M., Curwen, R., Hodges, S., Newman, J., Steggles, P., Ward, A., Hopper, A.: Implementing a Sentient Computing System. IEEE Computer 34(8), 50–56 (2001)
10. Hallberg, J., Nilsson, M., Synnes, K.: Positioning with Bluetooth. In: ICT (2003)
11. Borriello, G., Liu, A., Offer, T., Palistrant, C., Sharp, R.: WALRUS: Wireless Acoustic Location with Room-Level Resolution using Ultrasound. In: Mobicom (2005)
12. Peng, C., Shen, G., Zhang, Y., Li, Y., Tan, K.: BeepBeep: A High Accuracy Acoustic Ranging System using COTS Mobile Devices. In: SenSys (2007)
13. Navizon Technical Paper (2007),  
[http://www.navizon.com/Navizon\\_wifi\\_gps\\_and\\_cell\\_tower\\_positioning.pdf](http://www.navizon.com/Navizon_wifi_gps_and_cell_tower_positioning.pdf) (cited January 11, 2009)

14. Bossler, J., Jensen, J., McMaster, R., Rizos, C.: *Manual of Geospatial Science and Technology*, 1st edn. Taylor & Francis (2002)
15. Ghilani, C., Wolf, P.: *Adjustment Computations: Spatial Data Analysis*, 4th edn. John Wiley & Sons, Inc. (2006)
16. Filonenko, V., Cullen, C., Carswell, J.D.: Investigating Ultrasonic Positioning on Mobile Phones. In: Proc. of International Conference on Indoor Positioning and Indoor Navigation (IPIN). IEEE Xplore (2010)