

EE3842 Bachelor afstudeerproject

Bachelor thesis

Indoor Localization using Signals of Opportunity

Tuesday 9th July, 2013

Michael van der Cammen, Ramin Barak, Heming Zhang

Acknowledgements

Our gratitude goes out to J.A. Martinez Castaneda, TU Delft and Sony.

J.A. Martinez Castaneda has provided a lot of support and always shared his knowledge and valuable insights during the course of the project.

TU Delft has provided the study area and other supports to the project.

Sony has supplied us with multiple smart phones, which enabled us to perform the necessary experiments.

Foreword

This thesis is a part of the bachelor graduation project of Electrical Engineering. The topic is "Indoor Localization using Signals of Opportunity". Two groups focused on this particular topic, one of which mainly looked at the hardware/software involved and the other at the theoretical derivation of the algorithm. This thesis is concerned with the algorithmical part.

Contents

1	Introduction	6
2	Program of Requirements	7
2.1	Requirements of usage	7
2.2	Requirements of the ecological situation of the system within a particular environment	7
2.3	Requirements concerning the system design	7
2.3.1	Characteristics of usage	7
2.3.2	Characteristics of the production and installations	7
2.3.3	Characteristics of liquidation	7
2.4	Requirements of development of the system	7
3	Problem Formulation	8
4	Related Research	10
4.1	Time Of Arrival	10
4.2	Time Difference Of Arrival	11
4.3	Angle of arrival	13
4.4	"Fingerprinting"	14
5	Design process and proposed solution	15
5.1	Discussion of previous methods	15
5.2	Proposed Algorithm	15
5.2.1	The TDOA Phase	15
5.2.2	The Fingerprinting Phase	18
6	Tests and Results	26
6.1	Results of TDOA	26
6.2	Results of Fingerprinting	28
7	Discussion	33
8	Conclusion	35
9	Future work	36
	References	37
	Appendix A TDOA	39
	Appendix B Fingerprinting	41

Summary

This thesis proposes an algorithm to solve the issue of indoor localization using a source, reference receiver and mobile receiver placed in a convex room. The goal is determination of the position of the mobile receiver. The proposed algorithm consists of two localization techniques known from scientific literature: Time Difference of Arrival(TDOA) and Fingerprinting. In this work the choice was made to focus on two following kinds of signals: white noise and human speech. The proposed algorithm has been tested by means of various variables and these experiments have led to the following conclusions: the TDOA part of the algorithm strongly depends on the synchronization between the receivers while the Fingerprinting performs in an optimal manner when a white noise signal is being used as a source. The Fingerprinting part of the proposed algorithm contains four different ways to compare between the audio signal of the database and the audio signal of a real world situation. It appears that the cross - correlation in time domain yields the best results.

1 Introduction

The problem of self determining your location within an indoor environment has been a topic of interest for many years [1, 2, 3, 4]. The localization technology is applicable in many different situations, varying from finding your way in a congress hall to providing assistance to the fire department elements inside a building on fire [5].

Localization itself, however, is not an entirely new concept. Outdoor localization is already possible via technologies like the Global Positioning System (GPS). In this case the satellites simultaneously transmit their position and the current time to a receiver on earth, which uses the received data and the difference of arrival of the signals to derive its location [6]. Unfortunately, the signals produced by the satellites are not reliable in indoor environments due to attenuation during the propagation. This weakening is mainly caused due to the absorption of the satellite signals by most materials [7]. Because of this, outdoor signals can not be blindly used in an indoor environment. Another problem of the indoor localization process is the multipath effect which refers to the signal reflections due to the presence of ceilings, walls and other physical obstacles within indoor environments. All transmitted signals are reflected many times which makes the task of distinguishing the target signal from its reflections rather challenging [1].

Due to the fact that it is not always possible to make use of outdoor sources, this work focuses on making use of sources within a particular indoor situation. The most interesting signals to consider in this case, which are also the signals that are used throughout this work, are the so called "signals of opportunity": signals which are present in a particular environment either due to natural causes (e.g. heat or magnetic waves), or human intervention (e.g. WIFI signals), which are not originally meant for localization purposes. This particular type of signals is highly appealing due to their frequent presence in daily situations and also due to the fact that the localization process can be executed without the need to generate any signals. Many attempts have been made to solve the indoor localization problem by making use of signals of opportunity [8, 9, 10, 11]. In our particular case the decision fell upon using acoustic signals of opportunity, the main reason being the following: a small timing inaccuracy would result into huge deviations in distances in case of using electromagnetic waves due to their much higher speed.

The aim of this work is to design a robust and reliable algorithm to deal with the problems caused by the effects of multipath, which is equivalent to reverberation in case of acoustic signals. This work proposes a method to perform indoor localization while making use of the presence of particular acoustic sources. The algorithm in question is a combination of the Fingerprinting and the Time Difference Of Arrival (TDOA) techniques. TDOA will be explained in detail together with Fingerprinting in Sec.4. The implementation of the algorithm will result in an application to be run on the smart-phones provided by Sony.

The structure of this thesis is the following. Sec. 2 contain the programme of requirements. In Sec. 3 the problem statement of the research will be defined. Sec. 4 makes a summary of known ways to deal with indoor localization. Sec. 5 contains a detailed analysis of the proposed algorithm. Sec. 6 describes the various simulations and results. In Sec. 7 a discussion of the results is provided as well as possible improvements. Conclusions are drawn in Sec. 8. The last section suggests the possible future work.

2 Program of Requirements

This thesis proposes a new algorithm to deal with the problem of indoor localization using signals of opportunity. This particular topic has been of many interest to the scientific society for many years, but unfortunately no satisfactory solution has been derived yet. This work tries to add to those many years of research and proposes some interesting ideas.

2.1 Requirements of usage

2.1.1 The algorithm must be able to localize a mobile phone within a convex room.

2.1.2 The algorithm must not make use of any other signals than those of opportunity.

2.1.3 The algorithm must be performed quickly on the mobile phone.

2.2 Requirements of the ecological situation of the system within a particular environment

2.2.1 The algorithm must not transmit the information of the user to other devices.

2.3 Requirements concerning the system design

2.3.1 Characteristics of usage

2.3.1.1 The algorithm must be applicable to acoustic signals.

2.3.1.2 The algorithm must be implementable by mobile phones.

2.3.1.3 The algorithm must be performed automatically right after the start.

2.3.2 Characteristics of the production and installations

2.3.2.1 No additional signal source is needed.

2.3.2.2 The placing of the reference receiver must not interfere with the room or building's casual activities.

2.3.3 Characteristics of liquidation

2.3.3.1 The reference receiver must be easily demountable.

2.3.3.2 The reference receiver must be recyclable.

2.4 Requirements of development of the system

2.4.1 The algorithm will be tested using Matlab.

3 Problem Formulation

In this section, the mathematical model of the problem in question and the notation used in this thesis are defined first.

Consider a convex room where it is assumed that there is always a Line-of-Sight channel between the source and the receivers. The reason for having this Line-of-Sight channel will be explained in Sec. 5. Without loss of generality and for simplicity, the geometry of a room in this work is approximated by a rectangular cuboid, which is the usual shape of a room [12]. The x , y and z axis of the Cartesian coordinate system are defined along the length, width and height direction respectively. Fig. 1 shows the geometry together with the Cartesian coordinate system. For a room with length L_x , width L_y and height L_z , the location of one point within this room is given by the vector $\mathbf{p} = [x, y, z]^T$, where $0 < x < L_x, 0 < y < L_y, 0 < z < L_z$.

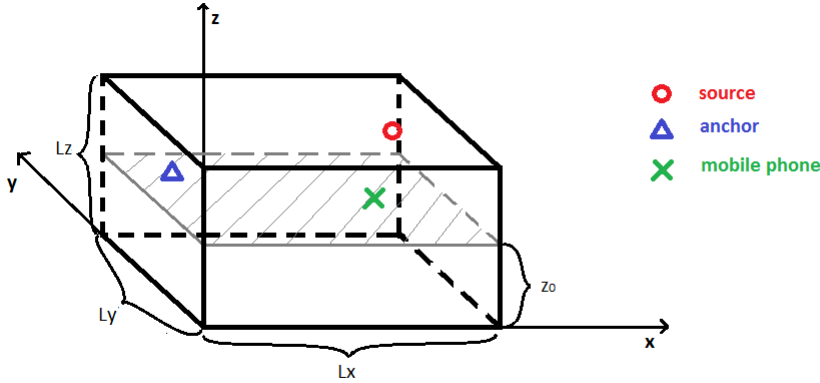


Figure 1: Scenario with 1 source, 1 anchor and 1 mobile receiver

The study of this work involves an acoustic signal source, a reference receiver (also referred to as an anchor, the case with more anchors will be discussed in Sec. 9), and a mobile receiver distributed around the room. The position of the source is defined as $\mathbf{p}_s = [x_s, y_s, z_s]^T$, the position of the anchor is defined as $\mathbf{p}_a = [x_a, y_a, z_a]^T$, and the position of the mobile receiver is defined as $\mathbf{p}_m = [x_m, y_m, z_m]^T$.

As described in Sec.1, signals of opportunity are the signals which are present in a particular environment and are originally not meant for localization purposes. Several examples of signals of opportunity are the WIFI signals from access points, the noises generated by all electronic devices and the announcements emanating from the loudspeakers. There are many different signals of opportunity, but all of them share the following property: the pattern of these signals is expected to be unknown in advance. For example, one is not to be expected to have any specific knowledge about what the noise will look like in advance. In this work, the type of the signals used is limited to acoustic signals. The main reason, as already mentioned in Sec.1, is the fact that the acoustic signals propagate much slower than electromagnetic waves. Therefore similar deviations in time measurements will result in less deviation during localization. Besides, the wave length of acoustic signals used is usually longer than that of electromagnetic waves. The reason for this statement will be explained in more detail in Sec.5.2.2. There were many different kinds of acoustic signals of opportunity to choose from, but this work focuses on the two following types: white noise and speech signals. There are two different reasons for this particular prerogative: the two selected signals are usually present in most environments and also appear to suit the implementation represented in this thesis. Both of these reasons will be dealt with in more detail in Sec.5.2.2.

The dimension of the room (L_x, L_y, L_z) , the position of the source \mathbf{p}_s and the location of the anchor \mathbf{p}_a are given. Since most individuals hold their mobile phone at a similar height, it is assumed that the height of the mobile phone is a known constant. In addition, the anchor is situated at the same height as the mobile phone. This height is being denoted as z_0 . Therefore the following is the case,

$$\begin{aligned}\mathbf{p}_a &= [x_a, y_a, z_a]^T \approx [x_a, y_a, z_0]^T, \\ \mathbf{p}_m &= [x_m, y_m, z_m]^T \approx [x_m, y_m, z_0]^T.\end{aligned}\tag{1}$$

Since z_0 is an estimation of the height, it introduces a certain inaccuracy. In Sec.5.2.2 it will be shown that this inaccuracy is acceptable.

Thus the task is to compute the x , y coordinates of the mobile receiver x_m and y_m by making use of the given information above.

Unlike out in the open, the propagation of waves indoors consists of many different paths due to the effect of reflection. A channel is defined as the set of all the propagation paths from a source to a receiver and is described with its impulse response h . Here a Wide-Sense-Stationary Uncorrelated Scattering(WSSUS) model is being used[13]. In this model, it is assumed that the channel correlation function is invariant over time and that the scatterers with different path delays are uncorrelated. In the discrete version, the impulse response is denoted as $h[n]$, or $H[\omega]$ in the frequency domain. Therefore, in the noiseless case, the signal received at the receiver can be written in the time and frequency domains using discrete Fourier transform as,

$$x[n] = s[n] * h[n] \xrightarrow{F} X[\omega] = S[\omega]H[\omega],\tag{2}$$

where $s[n]$ and $S[\omega]$ are the source signal in time and frequency domain, respectively.

Naturally, the noises can not be neglected in a real world scenario. There are several sources which contribute to the noise. One is the signal source which could be considered as a part of the source signals since the signals are unknown in advance. Another are the receivers and is uncorrelated with the source signals and the channel. Some noise might also be introduced during the propagation of the signal due to the changes in temperature, humidity or density of the medium. This particular kind of noise is also uncorrelated with the source signals and the channel. The change of signals due to multi-path effect, diffraction or other effects correlated with the channel is not considered as noise but a part of the channels. Therefore the noise could be split into the following two groups: part of the source signal and addition to the received signal. Taking all of the above mentioned noise into account, the following model arises,

$$x[n] = s[n] * h[n] + w[n] \xrightarrow{F} X[\omega] = S[\omega]H[\omega] + W[\omega],\tag{3}$$

where $w[n]$ and $W[\omega]$ are the noise in time and frequency domain, respectively. In this case it is assumed that the noise is uncorrelated to the source signals and the channel and varies per measurement. A common assumption of noise is white Gaussian noise [14, 15, 16]. In this case, the noise in frequency domain is also white Gaussian noise according to its property.

Furthermore, the convolution between sequences X and Y is denoted as $X * Y$ and the cross-correlation between sequences X and Y is denoted as $x * y$. The complex conjugate of x is denoted as x^* while the expected value of sequence X is denoted as $E[X]$. The covariance of sequences X and Y is denoted as $cov(X, Y)$, the standard deviation of a sequence X is denoted as σ_X and the 2-norm of vector \mathbf{x} is given by $\|\mathbf{x}\|$. The inverse of matrix \mathbf{X} is denoted as \mathbf{X}^{-1} while the velocity of the propagation of the sound wave in air is denoted as c . And finally the distance between the source and the anchor is denoted as $d_{s,a}$, the distance between the source and the mobile receiver is denoted as $d_{s,m}$ and the distance between the anchor and the mobile receiver is denoted as $d_{a,m}$.

4 Related Research

The concept of indoor localization has intrigued the scientific world throughout the past decades. It is therefore not surprising that many different approaches have been proposed to deal with the problem. In the following subsection a brief analysis is provided of the most prominent localization techniques relevant to this work.

4.1 Time Of Arrival

Time Of Arrival (TOA) is the simplest of these techniques. In this case the location of a particular source is being derived by means of a certain amount of receivers at known locations. Let us take a source at an unknown position P_s and a receiver at a known position P_{r1} . Let us also denote the time at which the signal is transmitted by the source as T_t and the time at which the signal is received by the receiver as T_{r1} . If both of these times are known, the distance between P_s and P_{r1} , denoted as D_{sr1} , may be derived using the following formula,

$$D_{sr1} = c(T_{r1} - T_t), \quad (4)$$

where c denotes the speed at which the signal propagates throughout the space and depends on the source used.

Fig. 2 shows an overview of the situation.

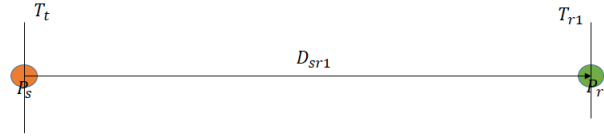


Figure 2: TOA basic situation

However, a single receiver with a known position does not suffice to derive the position of the unknown source. Two receivers with known positions, namely T_{r2} and T_{r3} , need to be added to the configuration in order to be able to use TOA to estimate the location of the source. Due to the fact that there are three receivers, there are three direct distances between the source and each one of the receivers, namely D_{sr1} , D_{sr2} and D_{sr3} . By making use of the Fig. 3 as well as Pythagoras' theorem the following equations can be derived,

$$(D_{sr1})^2 = (X_s)^2 + (Y_s)^2, \quad (5)$$

$$(D_{sr2})^2 = (X_2 - X_s)^2 + (Y_2 - Y_s)^2, \quad (6)$$

$$(D_{sr3})^2 = (X_3 - X_s)^2 + (Y_3 - Y_s)^2. \quad (7)$$

One way of solving (5), (6) and (7) and enabling the usage of more receivers (in case of inaccuracies or noise) is by means of the linear least-squares solution. This takes place in the following manner. Eq. (5) is extracted from (6) and also from (7) resulting into,

$$(D_{sr2})^2 - (D_{sr1})^2 = (X_2)^2 - 2X_2X_s + (Y_2)^2 - 2Y_2Y_s, \quad (8)$$

$$(D_{sr3})^2 - (D_{sr1})^2 = (X_3)^2 - 2X_3X_s + (Y_3)^2 - 2Y_3Y_s. \quad (9)$$

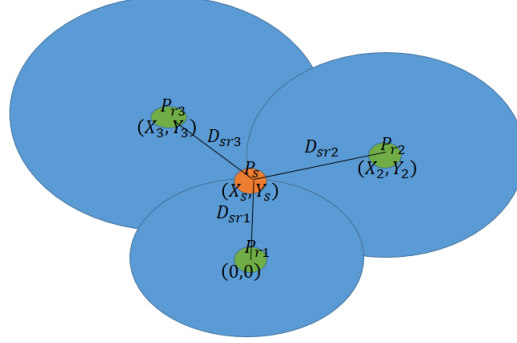


Figure 3: TOA with three receivers

A way of stating (8) and (9) in a matrix form is the following,

$$\mathbf{H}\mathbf{x} = \mathbf{b},$$

$$\text{where } \mathbf{H} \equiv \begin{bmatrix} X_2 & Y_2 \\ X_3 & Y_3 \end{bmatrix}, \mathbf{x} \equiv \begin{bmatrix} X_s \\ Y_s \end{bmatrix}, \mathbf{b} \equiv \frac{1}{2} \begin{bmatrix} (X_2)^2 + (Y_2)^2 - (D_{sr2})^2 + (D_{sr1})^2 \\ (X_3)^2 + (Y_3)^2 - (D_{sr3})^2 + (D_{sr1})^2 \end{bmatrix}. \quad (10)$$

In the case of (10) the linear least - squares solution would look like the following,

$$\hat{\mathbf{x}} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{b}. \quad (11)$$

The problem with the time of arrival method is synchronization. The absolute values of T_t , T_{r1} , T_{r2} and T_{r3} are difficult to derive due to different internal timing mechanisms. This timing problem results into inaccuracies during the calculation of P_s . Another obstacle in this case is the fact that while using signals of opportunity as sources, the signal transmission time instant might be completely unknown which makes the above calculations impossible. For more details on this technique the reader is referred to [5, 17, 18].

4.2 Time Difference Of Arrival

Time Difference Of Arrival (TDOA) is a method which makes use of the fact that a particular amount of receivers placed at different distances from a particular source receive the signals produced by the source at different times. Take three different receivers at known positions P_{r1} , P_{r2} and P_{r3} and a single source situated at an unknown position P_s . The whole situation is displayed in Fig. 4.

Assume the source transmits a particular signal towards each of the receivers. If the distances between the receivers from the source are denoted as D_{r1r2} and D_{r1r3} , then the following holds,

$$\sqrt{X^2 + Y^2 + Z^2} - \sqrt{(X - X_2)^2 + Y^2 + Z^2} = T_{r1r2}c = D_{sr1r2}, \quad (12)$$

$$\sqrt{X^2 + Y^2 + Z^2} - \sqrt{(X - X_3)^2 + (Y - Y_3)^2 + Z^2} = T_{r1r3}c = D_{sr1r3}, \quad (13)$$

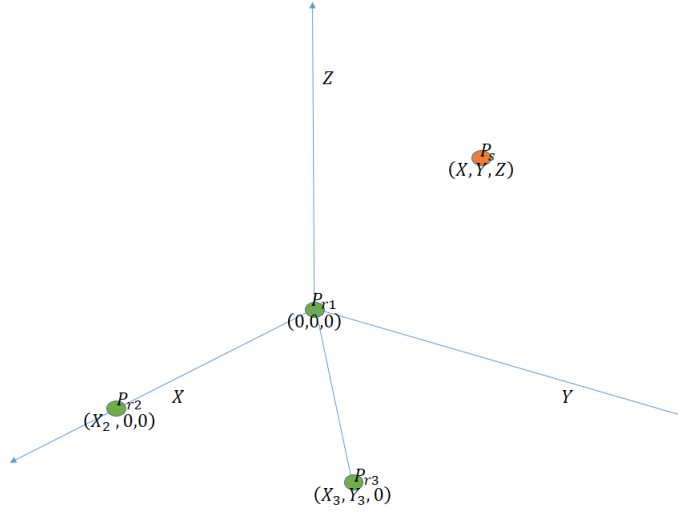


Figure 4: The initial position to derive TDOA comprising three receivers

where T_{r1r2} is the difference in time it takes for a signal to get from the source to the first receiver and from the source to the second receiver while T_{r1r3} represents the difference in time it takes for a signal to get from the source to the first receiver and from the source to the third receiver. c denotes the propagation speed of the signal and depends on the source being used.

The interesting part about (12) and (13) is the fact that D_{sr1r2} and D_{sr1r3} are fixed values which are caused by the fixed nature of T_{r1r2} (as long as neither the receivers nor the source are moved) and c . This actually enables us to consider the whole TDOA problem as the construction of hyperboloids. This particular switch is possible due to characteristic of the hyperboloids which states that the difference between the distances from a particular point to the loci of the hyperboloid is fixed. In this case the receivers would represent the loci of the hyperboloid and the source would represent a random point.

Rearranging, squaring and simplifying (12) and (13) results into the following,

$$(D_{sr1r2})^2 - X_2^2 + 2X_2X = 2D_{sr1r2}\sqrt{X^2 + Y^2 + Z^2}, \quad (14)$$

$$(D_{sr1r3})^2 - a^2 + 2X_3X + 2Y_3Y = 2D_{sr1r3}\sqrt{X^2 + Y^2 + Z^2}, \quad (15)$$

where a is equal to $\sqrt{(X_3)^2 + (Y_3)^2}$.

Squaring (14) and (15) provides us with the equations for two distinct hyperboloids, the loci of one of which are situated at the origin and $(X_2, 0, 0)$ and the other at the origin and $(X_3, Y_3, 0)$. Naturally, these two hyperboloids will intersect at a curve. However, in order to be able to estimate P_s another hyperboloid is required which would imply the addition of another receiver to the situation. However, it is also possible to turn the situation around: set off with an unknown receiver and four known sources and estimate the location of the receiver.

Although the TDOA is a rather attractive technique, there is one challenging problem: the synchronization between the different receivers needs to be extremely accurate to minimize localization errors. For more details the reader is referred to [3, 5, 8, 19, 20, 21, 22].

4.3 Angle of arrival

In the Angle of arrival (AOA) sometimes referred to as DOA (Distance of Arrival) technique the angle between the transmitted signal and some direction used as a reference is being derived. Consider a scenario which involves a source located at an unknown position (X_s, Y_s) and two receivers located at $(0,0)$ and (X_{r2}, Y_{r2}) as is pictured in Fig. 5. Using (5) as a reference the following can be derived for

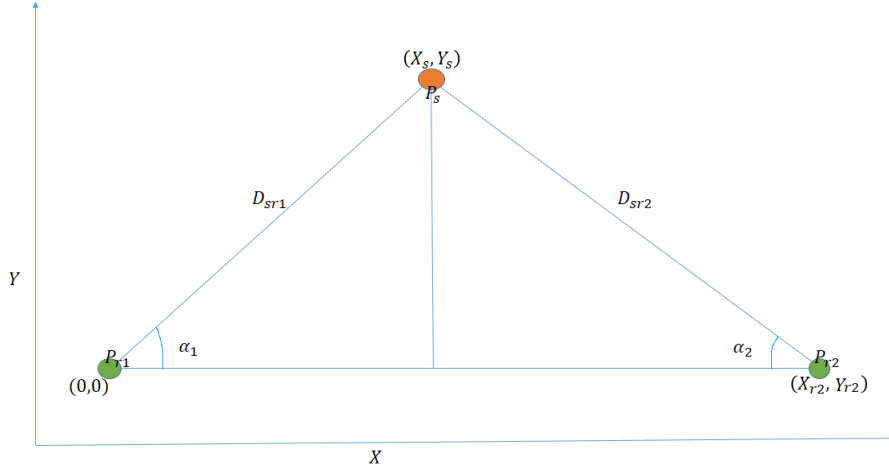


Figure 5: Depiction of a possible AOA scenario

the position of the source,

$$\begin{bmatrix} X_s \\ Y_s \end{bmatrix} = \begin{bmatrix} D_{sr1} \cos \alpha_1 \\ D_{sr1} \sin \alpha_1 \end{bmatrix}, \quad (16)$$

$$\begin{bmatrix} X_s \\ Y_s \end{bmatrix} = \begin{bmatrix} X_{r2} - D_{sr2} \cos \alpha_2 \\ Y_{r2} + D_{sr2} \sin \alpha_2 \end{bmatrix}. \quad (17)$$

Transforming (16) and (17) yields,

$$\mathbf{H}\mathbf{x} = \mathbf{b}, \quad (18)$$

where \mathbf{x} stands for $\begin{bmatrix} X_s \\ Y_s \end{bmatrix}$ and \mathbf{b} equals $\begin{bmatrix} D_{sr1} \cos \alpha_1 \\ D_{sr1} \sin \alpha_1 \end{bmatrix}$ in the first case. In the second case \mathbf{b} equals $\begin{bmatrix} X_{r2} - D_{sr2} \cos \alpha_2 \\ Y_{r2} + D_{sr2} \sin \alpha_2 \end{bmatrix}$.

The solution to the estimation of the source then lies in the linear least-squares methodology and is calculated in the following manner,

$$\hat{\mathbf{x}} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{b}. \quad (19)$$

The advantage of this particular technique is the fact that no synchronization for either the source or the receivers is necessary [5].

4.4 "Fingerprinting"

Fingerprinting involves the creation of a database of a particular space which will later serve as a site for indoor localization. The process can be conveyed in the following manner: a particular room, rooms or even an entire building is chosen where the technique of indoor localization will be applied. A particular source producing a signal will be placed within the environment and the room itself will be divided into evenly spaced grid points. Next, measurements of the received signals at each of these grid points will be performed and used to come up with a database of the room in question. This particular database may consist of grid points containing the whole received signal or only particular characteristics of it such as Received Signal Strength (RSS). The big advantage of this particular method is the uniqueness of the contents of each of the grid points enabling one to perform some interesting analysis. The disadvantages of fingerprinting are the large amount of time and effort required to construct the database as well as the spatial dependency. The latter implies the fact that any tiny alterations in the room in question would require a restart of the measurements [1, 23, 24].

Naturally, there are other existing approaches which enable indoor localization but the most of them are not related to this work.

5 Design process and proposed solution

5.1 Discussion of previous methods

Consider a room as described in Sec.3 Fig.1 containing a source at a known location, a reference receiver at a known position and a mobile receiver at an unknown position. The goal is to determine the position of the mobile receiver. First of all the existing methods from the scientific literature will be discussed.

- TOA makes use of the time at which a signal is being transmitted, T_t , as mentioned in Sec.4. Unfortunately, T_t is not known in our case and it is therefore not possible to make use of the TOA technique directly.
- The AOA (DOA) method determines the angle of arrival in order to calculate the location of the source. However, this method requires more receivers to determine this angle, which does not match the set up of this paper.
- The TDOA technique as described in Sec.4 can be used to determine the position of the mobile receiver in the room \mathbf{p}_m . Usually each TDOA measurement defines a hyperbola (or hyperboloid in case of 3D) and more than one TDOA measurement is needed to compute the intersection of the hyperbolas (hyperboloids in case of 3D). However, in order to do this, at least two additional sources (at known positions) are needed. Besides, the accuracy of the TDOA method is rather limited in case of an indoor environment.
- Fingerprinting is a relatively new method. The concept is to compare the received signal with a database constructed beforehand. This particular method does not try to eliminate the multipath effect but instead tends to make use of it. The most current researches on Fingerprinting comprise the storage of particular characteristics of the signals such as Received Signal Strength (RSS) in the database. The problem with this particular technique are its spatial dependency and effort involved in the production of the database.

5.2 Proposed Algorithm

The proposed algorithm is inspired by the TDOA and Fingerprinting methods. It is a combination of the TDOA measurements as well as the concept of Fingerprinting. Differences between the suggest technique and the existing methodology are being described in great detail in Sec. 5.2.1 and Sec. 5.2.2. TDOA measurements are employed to reduce the computational complexity and time consumption of the process by obtaining a region where the mobile receiver is located. The Fingerprinting is used for the more accurate localization. A brief block diagram is shown in Fig. 6. During the first phase, the TDOA between the receivers is extracted. Afterwards, the TDOA as well as the positions of the source, the reference receiver and the dimension of the room are used to obtain the possible region of situation of the mobile receiver. During the second phase, the received signals and the signals which make up the database are processed and compared to acquire the position of the mobile receiver.

5.2.1 The TDOA Phase

The TDOA Phase consists of two main stages:

1. Extracting the TDOA

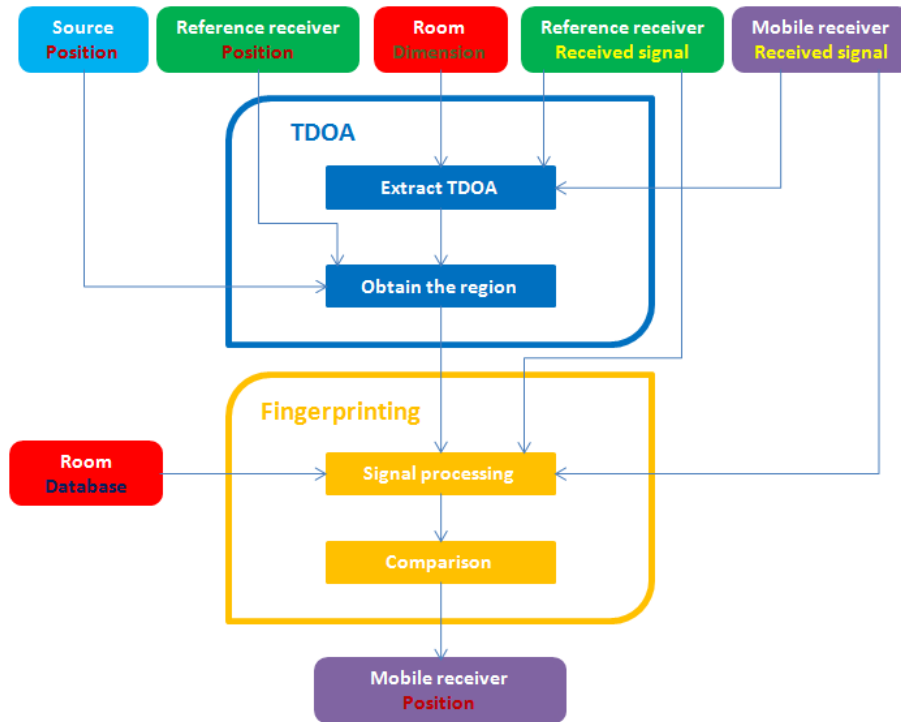


Figure 6: Block diagram algorithm

2. Determining the region of the target position

Stage 1. Extracting the TDOA

Consider two signals received by the reference receiver and the mobile receiver. If the distances from the source to the receivers are not same, then in the noiseless case, one signal could be considered as the other with some shift in the time domain. In the noise-present case, since the noises are uncorrelated as described in Sec. 3, the noise in one signal can be considered as a part of the signal. The other signal could be viewed as the time shift of the former signal with some additive noise.

Mathematically, these digital signals can be viewed as two sequences of samples and one sequence is considered as the superposition of the shift of the other sequence and some noise. This shift represents the TDOA between the two signals. In order to find the measure of this shift, one sequence is shifted again to find the best match with the other signal. The degree of linear coherence between these two sequences can be described by the correlation coefficient. The correlation of two vectors describes the degree of a linear relation between the two sequences. A correlation coefficient of 1 stands for a perfect positive linear relationship, a correlation coefficient of -1 means a perfect negative linear relationship and a correlation of 0 means no relationship at all. Higher correlation coefficient illustrates higher linear coherence.

The correlation coefficient of two sequences X and Y is computed as follows,

$$\begin{aligned}
r &= \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} \\
&= \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}.
\end{aligned} \tag{20}$$

This implies that the process of determining the measure of the shift will proceed as follows: one of the sequences is shifted and the correlation coefficient between this and the unshifted sequence is being computed. Naturally, the goal is to find the shift at which the correlation coefficient displays the highest value.

In case of long signals the shifting and comparing processes could result into rather high computational complexity. This complexity might be reduced by finding the maximum possible shift from some extra information such as the dimension of the room and the positions of the reference receiver and source. Given these information, the maximum possible TDOA can be derived. Assuming that the maximum possible distance from the source to a receiver is d_{max} and the distance between the source to the reference receiver is d_{sa} , then the maximum possible TDOA can be computed as:

$$TDOA_{max} = \frac{1}{c} \max\{d_{max} - d_{sa}, d_{sa}\}. \tag{21}$$

Therefore, the maximum measure of shift is derived:

$$sampleshift_{max} = \frac{F_s}{c} \max\{d_{max} - d_{sa}, d_{sa}\}, \tag{22}$$

where F_s is the sample rate of the signals.

In the practice, two receivers can not be perfectly synchronized, which means that the signals received by two receivers are also not perfectly synchronized. The consequence of this bad synchronization is that the obtained TDOA deviates from the real TDOA. Besides the synchronization, multi-path effect also causes inaccuracy. Due to the reflection and scattering, the received signal is a superposition of the signals from all the paths. As indicated in Sec. 3, it can be seen as the source signal convolved with the channel impulse response. Therefore the correlation between two received signals may not give an accurate TDOA. At last, the noise also increases the inaccuracy of obtained TDOA. The obtained TDOA is thus the sum of the real TDOA with the error due to synchronization, the error due to multi-path effect and the error due to the noise.

Stage 2. Determining the region

Without any synchronization error, noise, or multi-path effect, the obtained TDOA is equal to the real TDOA. Since the positions of the source and the reference receiver are given, the distance from the source to the reference receiver d_{sa} can be computed. From obtained TDOA, the difference between d_{sa} and the distance from the source to the mobile receiver d_{sm} can be derived. Therefore, d_{sm} can be obtained:

$$d_{sm} = d_{sa} + \Delta d \tag{23}$$

$$= d_{sa} + c \times TDOA \tag{24}$$

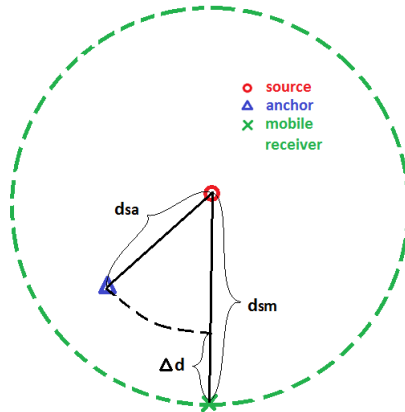


Figure 7: Sketch of TDOA

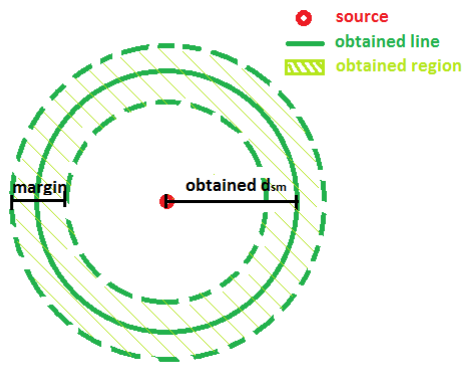


Figure 8: Sketch of obtained region from TDOA

Using d_{sm} , a circle around the source is obtained where the mobile receiver is located. A sketch is shown in Fig. 7.

In practice, the synchronization error, noise and multi-path effect are always present. Hence, the obtained TDOA has certain deviation from the real TDOA. This inaccuracy becomes a margin as shown in Fig. 8. The obtained region is thus a ring with the source in the center.

5.2.2 The Fingerprinting Phase

The Fingerprinting Part consists of three main stages:

1. Building the database
2. Signal processing
3. Comparing the received signals with the database

Stage 1. Building the database

Current Fingerprinting algorithms usually use some characteristics of the channel such as Received Signal Strength, Angle of Arrival or Time of Arrival to build the fingerprinting database

[1, 23, 24]. This implies that even small amendments to the rooms would require a complete reconstruction of the database. Although the concept of Fingerprinting is the same, the proposed implementation in this work is quite different. The real wave form samples are stored in the proposed database, which means that the database consists of the whole channels. This requires more memory than the conventional Fingerprinting, but provides more information about the room.

First of all an imaginary grid within the room is made. As mentioned in the previous section, a Cartesian coordinate system is defined for the room. The produced imaginary grid corresponds to this coordinate system and its axis are parallel to it. Due to this the coordinates of each point on the grid may be computed rather quickly as the location of that particular point. The Fingerprinting database consists of the measurements taken within the room in question. These particular measurements consist of audio samples which are recorded at each of the imaginary grid points. Each unit in the database consists of the location of a point and the corresponding audio sample.

Stage 2. Signal processing

Stage 2a. Filtering and down-sampling

It is important to consider the fact that there could be many small objects in the room. Each of them could reflect the sound waves, which contributes to the multi-path effect. The change of the location of these small objects would also change the channels. Since these small objects could be easily moved something needs to be done to reduce the impact of the reflections introduced by them. Due to this signals of long wave lengths are preferred. For a sound wave, a phase difference between two points is considered stable if the change of the phase difference for a single frequency is small. In order to have a stable phase difference between two grid points, a long wave length is expected compared with the distance between two neighbouring grid points. If the wave length is much smaller than the distance between two neighbouring grid points, the phase difference of the wave between grid points can vary a lot. Hence, long wave lengths, or in other words, low frequency components are desired. In a room with strong multi-path effect, the channel can be considered correlated only over a distance of $\lambda/2$, where λ is the wave length [5]. Then an adequate distance d between two neighbouring grid points is about λ . Because Fingerprinting can not cover all the positions in a room. Once the mobile receiver is located between two neighbouring grid points, we would like the channel from the source to the mobile receiver to be correlated to the channel from the source to a grid point nearby the mobile receiver. If d is much longer than λ , then the channel from the source to the mobile receiver can be uncorrelated to all channels to the nearby grid points.

The most common sampling rates for digital audio are 44.1 kHz, 48 kHz, 88.2 kHz, 96 kHz and 192 kHz [25]. Even for the lowest sampling rate 44.1 kHz, the wave length is less than 1 cm. The sampling rate is thus much higher than the desired frequency components. Usually it is possible to change the sampling rate of the mobile phone with an anti-aliasing filter. If this is not the case both down-sampling and filtering has to be performed.

There are many different low-pass filters, among which the Butterworth filter was chosen [26]. The most important property of a Butterworth filter is its possession of a maximally flat magnitude in the passband. The Butterworth filter is chosen to implemented because

the distortion to low frequency component is minimal. The cut-off frequency of the filter is determined by the distance between grid points and the original audio sampling rates. For example, if the distance between nearby grid points is around 0.5 m, for a sound wave with similar wave length, the frequency is around 680 Hz. The Nyquist frequency will be around 1360 Hz in this particular case. If the original sampling frequency is 44.1 kHz, then 1470 Hz = 44100 Hz/30 is a good choice. In this case an integer ratio between the original sampling frequency and the cut-off frequency is desired due to its convenience for down-sampling. As soon as filtering has taken place, the highest frequency component is much lower than the original sampling frequency, which means that the signals can be down-sampled without losing information. The down-sampling could reduce the amount of computing work and is therefore a crucial process. In the digital world, signals are some sequences. If an arbitrary ratio is chosen for the down-sampling, interpolation of the signals needs to take place before the implementation of the down - sampling. Naturally, this would increase the computational complexity. An integer ratio is thus desired.

Stage 2b. Processing the database and the received signals

Since use is being made of signals of opportunity, there is no knowledge about the received signals which will be compared later on. Therefore the signal to be compared can be considered as a different signal from the database. Before the comparison, both the received signals and the signals in the database need to be processed to eliminate the impact of the different source signals.

As defined in the previous section, a mathematical model could be constructed for this particular problem. In this particular case two different source signals are being taken into consideration. One is the source signal used to build the database, denoted as s_A . And the other is the signal sent by the source when localizing the mobile phone, which is denoted as s_B . Since there are two microphones on the smart-phone, two different channels are being considered. The impulse response of the channel from the source to one microphone is defined as h_1 and the impulse response from the source to the other microphone is defined as h_2 . First the noiseless case will be looked at. One point in the database consists of two sequences defined in time and frequency domains as follows,

$$\begin{aligned} x1[n] = s_A[n] * h_1[n] & \xrightarrow{F} X1[\omega] = S_A[\omega]H_1[\omega] \\ x2[n] = s_A[n] * h_2[n] & \xrightarrow{F} X2[\omega] = S_A[\omega]H_2[\omega]. \end{aligned} \quad (25)$$

The received signals at the same point can be defined in similar way,

$$\begin{aligned} x3[n] = s_B[n] * h_1[n] & \xrightarrow{F} X3[\omega] = S_B[\omega]H_1[\omega] \\ x4[n] = s_B[n] * h_2[n] & \xrightarrow{F} X4[\omega] = S_B[\omega]H_2[\omega]. \end{aligned} \quad (26)$$

As described in Sec.3, this thesis focuses on two kind of source signals, white noise and speech signal. These two are common acoustic signals which can be experienced in a room rather easily. Possible occurrences are the noise from an air conditioner or a speaker used for announcements. For example, white noise from an air conditioner or a loudspeaker for announcement. White noise is stationary, which implies that its joint probability distribution does not change with the shift in time or space. While speech signal is less ideal than white noise: it is not stationary, but it has only short-term correlations [27]. Therefore, it can be assumed that it is not correlated in the long term.

There are four proposed methods for the signal processing.

- Cross-correlation

First of all the cross-correlation between the signals received at two microphones is computed.

In time domain, the cross-correlation between x_1 and x_2 can be written as

$$\begin{aligned}
 (x_1 \star x_2)[n] &= x_1[-n] \star x_2[n] = (s_A[-n] \star h_1[-n]) \star (s_A[n] \star h_2[n]) \\
 &= (s_A[-n] \star s_A[n]) \star (h_1[-n] \star h_2[n]) \quad (27) \\
 &= r_A[n] \star (h_1[-n] \star h_2[n]).
 \end{aligned}$$

The sequence $r_A[n]$ is the auto-correlation sequence of $s_A[n]$. For a long stationary sequence $s_A[n]$, $r_A[n]$ is expected to have a pronounced peak at $r_A[0]$ with relatively small side-lobes. White noise is a good example of this phenomenon. Speech signal is not the best case, but usually a pronounced peak can be expected. Therefore (27) can be approximately reduced to

$$(x_1 \star x_2)[n] \approx r_A[0](h_1[-n] \star h_2[n]). \quad (28)$$

Similarly, we get

$$(x_3 \star x_4)[n] \approx r_B[0](h_1[-n] \star h_2[n]). \quad (29)$$

$r_B[n]$ is the auto-correlation sequence of $s_B[n]$. Therefore, the comparison between the database and the received signals becomes the comparison between the results of (28) and (29). However, if the side-lobes of $r_A[n]$ or $r_B[n]$ are not small enough, then due to the convolution in (27), the side-lobes will be convolved with $(h_1[-n] \star h_2[n])$. Naturally, this will also contribute to the cross-correlation of two signals and might possible result into a miss match between the received signal and the database.

A similar operation could be performed in frequency domain and would look like the following,

$$\begin{aligned}
 (X_1 \star X_2)[\omega] &= \sum_{\omega_0=-\infty}^{\infty} X_1^*[\omega_0]X_2[\omega + \omega_0] = \sum_{\omega_0=-\infty}^{\infty} S_A^*[\omega_0]H_1^*[\omega_0]S_A[\omega + \omega_0]H_2[\omega + \omega_0] \\
 &= \sum_{\omega_0=-\infty}^{\infty} S_A^*[\omega_0]S_A[\omega + \omega_0]H_1^*[\omega_0]H_2[\omega + \omega_0]. \quad (30)
 \end{aligned}$$

If the source signal $S_A[\omega]$ is wide-sense stationary, then the autocorrelation of S_A is defined as $R_{SS_A}[\omega] = E[S_A[\omega + \omega_0]S_A^*[\omega_0]]$. For certain channels H_1 and H_2 , with wide-sense stationary random process $S_A[\omega]$, the expected value of $(X_1 \star X_2)[\omega]$ is

$$\begin{aligned}
 E[(X_1 \star X_2)[\omega]] &= E\left[\sum_{\omega_0=-\infty}^{\infty} S_A^*[\omega_0]S_A[\omega + \omega_0]H_1^*[\omega_0]H_2[\omega + \omega_0]\right] \\
 &= \sum_{\omega_0=-\infty}^{\infty} E[S_A^*[\omega_0]S_A[\omega + \omega_0]]H_1^*[\omega_0]H_2[\omega + \omega_0] \quad (31) \\
 &= R_{SS_A}[\omega] \sum_{\omega_0=-\infty}^{\infty} H_1^*[\omega_0]H_2[\omega + \omega_0] \\
 &= R_{SS_A}[\omega](H_1 \star H_2)[\omega].
 \end{aligned}$$

Although it is known that the expected value of a sequence is not equal to the sequence itself, in case of the standard deviation of $S_A[\omega]$ being relative small, it is expected that the expected value is close to the real value. Eq. (30) can be thus approximately reduced to

$$(X_1 \star X_2)[\omega] \approx R_{SS_A}[\omega](H_1 \star H_2)[\omega]. \quad (32)$$

And in the same way, we get

$$(X_3 \star X_4)[\omega] \approx R_{SS_B}[\omega](H_1 \star H_2)[\omega]. \quad (33)$$

White noise is stationary in time as well as frequency domain. Thus this method can only be applied to source signals with small standard deviation as explained above. Thus the application is very limited.

- Source cancellation

As described in (25) and (26), the time domain signals in the database and received by the mobile phone could be described as the convolution of the source signals and the channels. The convolution of the signals are computed as follows,

$$\begin{aligned} x_1[n] \star x_4[n] &= (s_A[n] \star h_1[n]) \star (s_B[n] \star h_2[n]) = s_A[n] \star s_B[n] \star h_1[n] \star h_2[n], \\ x_2[n] \star x_3[n] &= (s_A[n] \star h_2[n]) \star (s_B[n] \star h_1[n]) = s_A[n] \star s_B[n] \star h_1[n] \star h_2[n]. \end{aligned} \quad (34)$$

Using this method, the impact of the different sources is eliminated. However, if there is some time difference between the signals, for example the random processing delay of the mobile phone, then after computing the convolution, the difference between two sequences could be large.

Another method is to cancel the sources in frequency domain. From (25) and (26), we have

$$\frac{X_1[\omega]}{X_2[\omega]} = \frac{S_A[\omega]H_1[\omega]}{S_A[\omega]H_2[\omega]} = \frac{H_1[\omega]}{H_2[\omega]} \quad (35)$$

$$\frac{X_3[\omega]}{X_4[\omega]} = \frac{S_B[\omega]H_1[\omega]}{S_B[\omega]H_2[\omega]} = \frac{H_1[\omega]}{H_2[\omega]}. \quad (36)$$

In the frequency domain, delay of several samples does not introduce a large error in the time domain. Unfortunately, this particular method has some disadvantages as well. If some frequency components of $h_2[n]$ are relatively small, the quotient at these frequencies will be quite large, possibly close to infinity. This problem could be solved by neglecting some frequency components. However, such a procedure varies per point and would result into extra computing work.

- Source comparison

In contrast to the method Source cancellation, the sources could also be compared instead. In time domain, a deconvolution would have to be performed twice. First the channel $h_1[n]$ and $h_2[n]$ are computed from deconvolution of $x_1[n]$, $s_A[n]$ and $x_2[n]$, $s_A[n]$. Afterwards, $s_B[n]$ is computed from deconvolution of $x_3[n]$, $h_1[n]$ and $x_4[n]$, $h_2[n]$ respectively. If the location of the mobile phone matches with a particular grid point of the database, the computed $s_B[n]$ should be similar. However, deconvolution is not a desired method. Although many deconvolution algorithms exist, they require a lot of computation and the results could be unstable. Besides, the time difference

between the signals also introduces more error. Therefore, another approach is used. Similar to the method of source cancellation, it is also processed in frequency domain,

$$\frac{X_3[\omega]}{X_1[\omega]} = \frac{S_B[\omega]H_1[\omega]}{S_A[\omega]H_1[\omega]} = \frac{S_B[\omega]}{S_A[\omega]} \quad (37)$$

$$\frac{X_4[\omega]}{X_2[\omega]} = \frac{S_B[\omega]H_2[\omega]}{S_A[\omega]H_2[\omega]} = \frac{S_B[\omega]}{S_A[\omega]}. \quad (38)$$

This method does not only solve the problem concerning deconvolution and time delay, but also has an advantage in case of the division. Since $s_A[n]$ is the source signal used to build the database it is possible to choose some broadband signals such as white noise as source signals intentionally. Due to this the quotients are much more stable than in the case of the method of source cancellation.

According to the statements above, experiments were carried out using four methods: cross-correlation both in time and frequency domains, source cancellation in frequency domain and source comparison in frequency domain. The accuracy yielded by these methods is discussed in Sec. 6.

Stage 3. Comparing the received signals with the database

In previous stage, a theoretical model without noise was used to derive the algorithms. If there is entirely no noise, the comparison between the database and received signals is simple, since it can be expected that the best fit is between two of the same signals. In practice, however, there is always noise. Therefore an analysis concerning the impact of the noise has been performed in order to choose the best technique of comparison. Using (3), the model of signals become,

$$\begin{aligned} x_1[n] &= s_A[n] * h_1[n] + w_1[n], \\ x_2[n] &= s_A[n] * h_2[n] + w_2[n], \\ x_3[n] &= s_B[n] * h_1[n] + w_3[n], \\ x_4[n] &= s_B[n] * h_2[n] + w_4[n]. \end{aligned} \quad (39)$$

Eq. (27) of the Cross-correlation method in the time domain becomes,

$$\begin{aligned} (x_1 \star x_2)[n] &= (s_A[-n] * h_1[-n] + w_1[n]) * (s_A[n] * h_2[n] + w_2[n]) \\ &= (s_A[-n] * h_1[-n]) * (s_A[n] * h_2[n]) + x_2[n] * w_1[n] + (s_A[-n] * h_1[-n]) * w_2[n] \\ &\approx r_A[0](h_1[-n] * h_2[n]) + x_2[n] * w_1[n] + (s_A[-n] * h_1[-n]) * w_2[n] \\ &\approx r_A[0](h_1[-n] * h_2[n]) + w_{CC-T}[n]. \end{aligned} \quad (40)$$

Compared with (27), the noises are convolved by x_2 and $s_A[-n] * h_1[-n]$. The result could be considered as adding a new noise $w_{CC-T}[n]$.

Eq. (30) of the Cross-correlation method in the frequency domain becomes,

$$\begin{aligned}
(X_1 \star X_2)[\omega] &= \sum_{\omega_0=-\infty}^{\infty} (S_A^*[\omega_0]H_1^*[\omega_0] + W_1^*[\omega_0])(S_A[\omega + \omega_0]H_2[\omega + \omega_0] + W_2[\omega + \omega_0]) \\
&= \sum_{\omega_0=-\infty}^{\infty} S_A^*[\omega_0]H_1^*[\omega_0]S_A[\omega + \omega_0]H_2[\omega + \omega_0] \\
&\quad + \sum_{\omega_0=-\infty}^{\infty} (X_2[\omega + \omega_0]W_1^*[\omega_0] + S_A^*[\omega_0]H_2^*[\omega_0]W_2[\omega + \omega_0]) \\
&= \sum_{\omega_0=-\infty}^{\infty} S_A^*[\omega_0]H_1^*[\omega_0]S_A[\omega + \omega_0]H_2[\omega + \omega_0] + W_{CC-F}[\omega].
\end{aligned} \tag{41}$$

Similar to the time domain, the noises are convolved by X_2 and $S_A H_2$. The result could be viewed as adding a new noise $W_{CC-F}[\omega]$.

Eq. (36) of source cancellation method becomes,

$$\begin{aligned}
\frac{X_1[\omega]}{X_2[\omega]} &= \frac{S_A[\omega]H_1[\omega] + W_1[\omega]}{S_A[\omega]H_2[\omega] + W_2[\omega]} \\
&= \frac{H_1[\omega] + \frac{W_1[\omega]}{S_A[\omega]}}{H_2[\omega] + \frac{W_2[\omega]}{S_A[\omega]}}.
\end{aligned} \tag{42}$$

Compared with (36), the noises are scaled by $S_A[\omega]$. Further on the impact of noise is scaled even more by the division.

Eq. (38) of the source comparison method becomes,

$$\begin{aligned}
\frac{X_3[\omega]}{X_1[\omega]} &= \frac{S_B[\omega]H_1[\omega] + W_3[\omega]}{S_A[\omega]H_1[\omega] + W_1[\omega]} \\
&= \frac{S_B[\omega] + \frac{W_3[\omega]}{H_1[\omega]}}{S_A[\omega] + \frac{W_1[\omega]}{H_1[\omega]}}.
\end{aligned} \tag{43}$$

Compared with (38), the noises are scaled by $H_1[\omega]$. Also, the impact of noise is scaled even further by the division.

As witnessed from the above mentioned analysis, it is clear that there are two kinds of noise which tend to have an impact on the proposed algorithm. One is the convolution of noise with the signals and the other is the scaling of noise by signals or channels.

There are many possible methods of comparison. The essence of comparison between the database and the received signals is comparison between two sequences of numbers, or two vectors as viewed from the perspective of linear algebra. The difference between two vectors needs to be defined by means of a measure. This measure needs to be robust against noise and linear transformations, and also save computing work. Robustness in case of linear translation implies that this measure between X, Y and $\alpha X + \beta, \gamma Y + \delta$ ($\alpha, \beta, \gamma, \delta$ are constants) are the same. This is rather useful in case of the suggested algorithm due to the fact that scaling and translation of the signals are common issues of the environment and the devices themselves. Several definitions of difference measure are listed as follows:

- Norm

There are different definitions of norms. The most common one is the so - called 2-norm. For example, consider $\mathbf{a} = [a_1, a_2, a_3] \in \mathcal{R}^3$, then $norm(\mathbf{a}) = \sqrt{a_1^2 + a_2^2 + a_3^2}$. If the dimension of the vectors is large, the norm is not an advisable choice. The problem with this particular method is the fact that too many vectors could have the same norm while some of them could still be unrelated at the same time. Thus this particular method is not robust in a noisy environment. Another disadvantage is being caused by the fact that the norm of a vector changes in case of a linear transformation.

- Projection

Another method to compare two vectors is by means of projecting one of the vectors onto the other and then compute the difference. In case of vectors \mathbf{x} and \mathbf{y} , the projection of \mathbf{x} on \mathbf{y} is,

$$Proj_{\mathbf{y}}(\mathbf{x}) = \left(\frac{\mathbf{x} \cdot \mathbf{y}}{\mathbf{y} \cdot \mathbf{y}} \right) \mathbf{y}. \quad (44)$$

And the relative difference of $Proj_{\mathbf{y}}(\mathbf{x})$ and \mathbf{y} is,

$$\frac{\|\mathbf{y} - Proj_{\mathbf{y}}(\mathbf{x})\|}{\|\mathbf{y}\|}. \quad (45)$$

This comparison technique does give a measure of the difference between two vectors. But then again, it does vary with linear transformations of the vectors.

- Correlation

The advantage of using correlation coefficient is caused by its invariance to linear transformations. This particular property is really useful in practice. One problem of the correlation coefficient is that if one of the sequences has all the same elements, the correlation coefficient will become undefined, since the variance of it will be zero. Theoretically, this could lead to a problem for the methods Source cancellation and Source comparison, if a quotient of two sequences contains all the same elements. However, in practice, due to the noise, the chance that a sequence is the product of a factor and another sequence is so small that it is neglectable. Besides, according to the model described in Sec. 3, different channels and different noises are uncorrelated. Moreover, as explained in our analysis above, the signals are also (almost) uncorrelated. Therefore, the correlation coefficient is a good measure to distinguish between different channels.

Taking all of the above mentioned information into consideration, correlation was chosen as the measure for the comparison. This due to its invariance to the linear transform, which is the desired property for the signals from different recorders and the fact that the theoretical disadvantage can be neglected in practice.

6 Tests and Results

The algorithm has been tested in the EEMCS building. Two mobile phones provided by Sony were used during the tests. In case of the TDOA test one mobile phone was used as an anchor and the other one as the mobile receiver. During the Fingerprinting tests both of them were used to build the database and to record received signals as well. Several tests have been performed for different purposes as follows:

1. This test had taken place in a room of the EEMCS building. A 3×10 grid was made within this rectangular room. The distance between two neighbouring grid points was around 0.5 m. Two mobile phones were used to test the TDOA algorithm. The synchronization of the mobile phones has been dealt with in [28]. At this time, the synchronization problem was not solved yet. Therefore the expectation was a relatively large deviation of time between two mobile phones: the results can be viewed in Sec 6.1.
2. In another room of the EEMCS building, another test concerning the TDOA part of the algorithm has been performed. Also this time two mobile phones were used during the test. The difference with the first test is the application of the manual synchronization which was expected to improve the accuracy of the experiment.
3. The fingerprinting algorithm was tested in the same room with the same settings as the first test. In this case white noise and speech signals were used as source signals. 8 sample were taken per point per source signal.

After the measurements the algorithm was implemented by means of Matlab which led to the results shown below.

6.1 Results of TDOA

Test 1

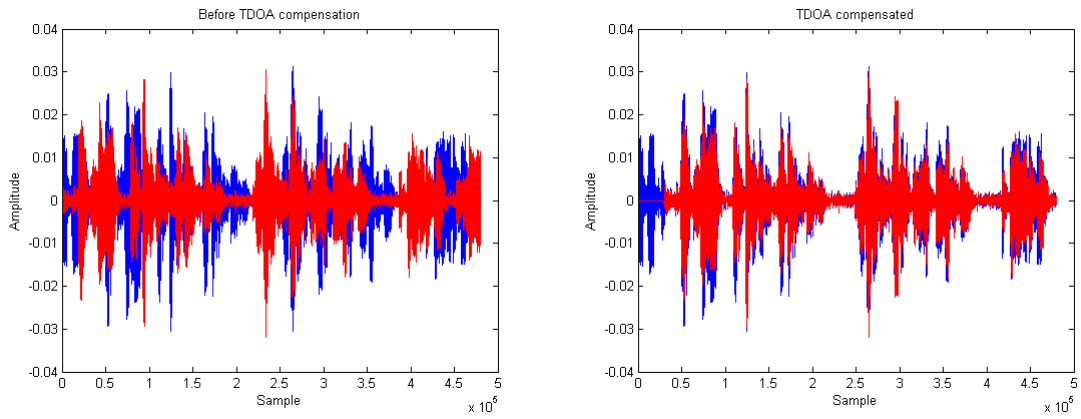
The process of determining the time shift between two signals is shown in Fig. 9. The originally received two signals are shown in Fig. 9(a). Fig. 9(c) illustrates the relation between the shift and the correlation coefficient. The shift corresponding to the best match can be derived from that particular figure. Applying this shift to the received signals, the fitted signals are plotted in Fig. 9(b).

The time difference of arrival originating from the measurement and processed by the algorithm is defined as $TDOA_{computed}$. The actual time difference of arrival is defined as $TDOA_{actual}$. The distance deviation Δd of $TDOA_{computed}$ and $TDOA_{actual}$ is defined,

$$\Delta d(i) = c(TDOA_{computed} - TDOA_{actual}), i = 1, 2, 3, \dots, N. \quad (46)$$

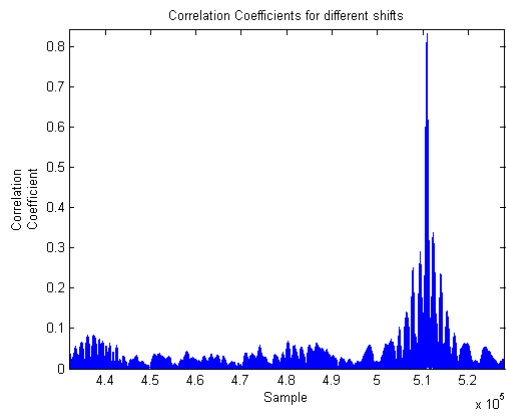
N is the number of pair of points where the measurements have been performed.

In Fig. 10, the distance deviation of $TDOA_{computed}$ and $TDOA_{actual}$ is plotted against the location of the grid points. This particular test measured the speech signal received at point (0, 1) against all grid points. The measurements of the remaining points against all grid points led to similar results, that is, the deviation is still large but seems to decrease as the distance between the mobile and reference receivers increases.



(a) Received two signals

(b) After shifting back



(c) Correlation coefficients

Figure 9: The process of determining the time shift between two signals

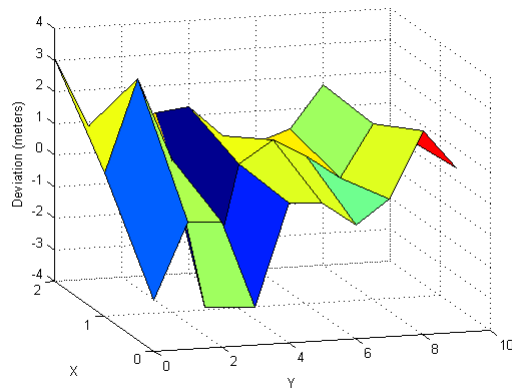


Figure 10: The distance deviation of TDOA method between point(0,1) and other points

Table 1: Results of test 2

Position	Measurement	$TDOA_{computed}(ms)$	$TDOA_{actual}(ms)$	difference (m)
1	1	11.00	8.74	-0.7744
1	2	11.07	8.74	-0.7978
1	3	10.79	8.74	-0.7044
2	1	3.92	4.22	0.1037
2	2	1.54	4.22	0.9208
2	3	3.83	4.22	0.1348
3	1	13.11	11.54	-0.5382
3	2	13.38	11.54	-0.6316
3	3	13.52	11.54	-0.6783
4	1	11.90	13.49	0.5443
4	2	11.00	13.49	0.8556
4	3	10.34	13.49	1.0813

The Root-mean-square-deviation(RMSD) of the difference of distance is computed as follows,

$$RMSD_d = \sqrt{\frac{\sum_{i=1}^N \Delta d^2(i)}{N}} = 2.5686m. \quad (47)$$

However, this particular result is not really accurate due to a lack of a reasonable method of synchronization at the time. The consequence is that some TDOA measurements display large deviation which is not caused by the TDOA algorithm itself. The details are discussed in Sec. 7.

Test 2

This test was performed in order to determine whether the TDOA part of the proposed algorithm would function in a real world situation. During the test the signals were synchronized manually. This was accomplished by transmitting a pulse (caused by a clap of hands) at an equal distance from the two receivers.

The reference receiver and the source were placed in such a way that the distance $d_{s,a}$ was equal to 5.96 meters. The mobile receiver was located at four different positions, resulting in $d_{s,m} = 2.96, 4.51, 2.0, 1.33$ meters. For each position three measurements were performed.

The results of the measurement and subsequent TDOA extraction are shown in Tab. 1. The $RMSD_d$ of this test is equal to 0.7045 m.

6.2 Results of Fingerprinting

Test 3

As described in the previous section, after using the TDOA algorithm, a small region is obtained where the mobile phone can be found. The Fingerprinting algorithm will be applied to this small region to improve the accuracy of the localization. Therefore, the accuracy of Fingerprinting also depends on the accuracy of the TDOA method. The smaller the region provided by the TDOA, the higher the accuracy of the output of the Fingerprinting part. In order to study the

accuracy of Fingerprinting given a certain region, the Fingerprinting algorithm was tested with different radius of the region. During the analysis the radius varied from $\sqrt{2}/2$ m to $5\sqrt{2}/2$ m. Here a step of $\sqrt{2}/2$ m is used for the radius.

There are 30 grid points, 2 source signals and 8 wave forms per point per signal. Using one wave form as database and one other wave form as received signal there is a total of 240 combinations for each region.

Four methods introduced in Sec. 5 were applied to all the grid points in the given region. As a reference, a random method was also computed. Firstly the correlation coefficient for each point is derived. During the random method, a uniform distributed random value as correlation coefficient for each point was generated. Furthermore, the error of the correlation coefficient for each combination of the signal types is defined as,

$$e_c(i, j) = \frac{C_{max}(i, j) - C_{receiver}(i, j)}{C_{max}(i, j)}, \quad i = 1, 2, 3, \dots, 30, \quad j = 1, 2, 3, \dots, 56, (57, \dots, 64). \quad (48)$$

where C_{max} is the maximum of correlation coefficient in this region, and $C_{receiver}$ is the correlation coefficient using j-th combination of wave forms at the i-th point, where the mobile receiver is situated. If the same type of signals are used for both received signals and the database, there are $7 \times 8 = 56$ combinations of wave forms. In other cases there are $8 \times 8 = 64$ combinations of wave forms. The total error of the correlation coefficients for all the grid points is defined as,

$$e_{c,total} = \frac{1}{30 \times 56} \sum_{i=1}^{30} \sum_{j=1}^{56} e_c(i, j), \quad \text{for same type of signals,} \quad (49)$$

$$e_{c,total} = \frac{1}{30 \times 64} \sum_{i=1}^{30} \sum_{j=1}^{64} e_c(i, j), \quad \text{for different type of signals.}$$

In the same way, the difference between the maximum correlation coefficient and the averaged correlation coefficient is defined as,

$$e_{g,total} = \frac{1}{30 \times 56} \sum_{i=1}^{30} \sum_{j=1}^{56} e_g(i, j), \quad \text{for same type of signals,}$$

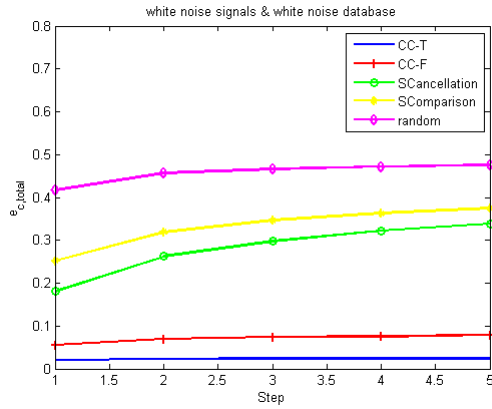
$$e_{g,total} = \frac{1}{30 \times 64} \sum_{i=1}^{30} \sum_{j=1}^{64} e_g(i, j), \quad \text{for different type of signals,}$$

where $e_g(i, j) = \frac{C_{max}(i, j) - C_{average}(i, j)}{C_{max}(i, j)}, \quad i = 1, 2, 3, \dots, 30, \quad j = 1, 2, 3, \dots, 56, (57, \dots, 64).$ (50)

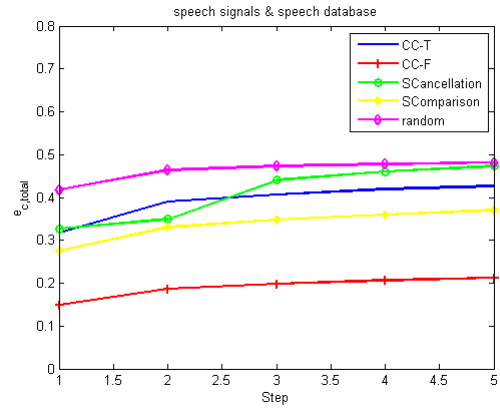
The results are shown in Fig. 11 and 12. These will be analysed in Sec. 7.

Afterwards, the distance between the estimated location and the actual location is computed. The distance deviation d of the j-th combination of wave forms at i-th point is,

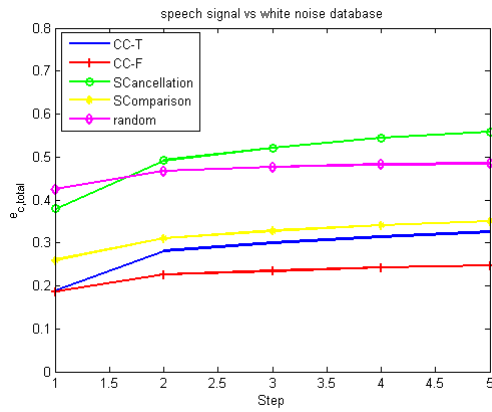
$$d(i, j) = \sqrt{(X_m(\tilde{i}, j) - X_m(i, j))^2 + (Y_m(\tilde{i}, j) - Y_m(i, j))^2}, \quad i = 1, 2, 3, \dots, 30, \quad j = 1, 2, 3, \dots, 56, (57, \dots, 64). \quad (51)$$



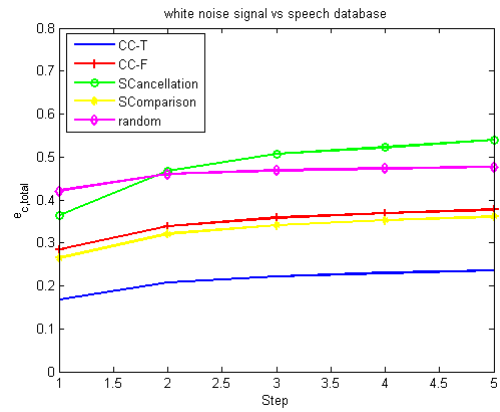
(a) white noise signals & white noise database



(b) speech signals & speech database

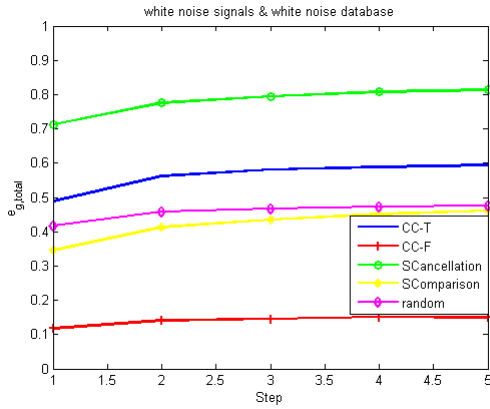


(c) speech signals & white noise database

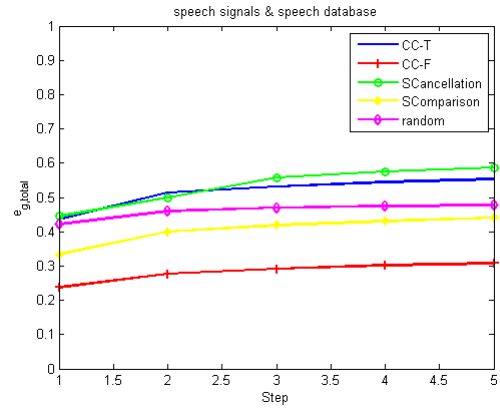


(d) white noise signals & speech database

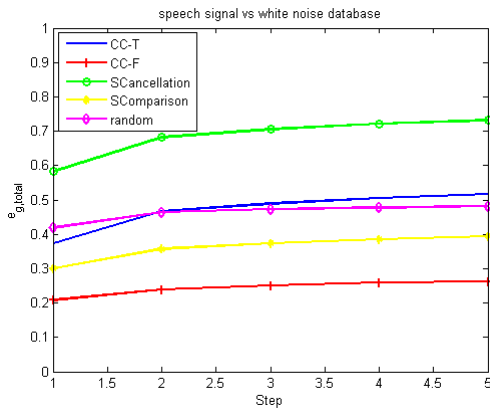
Figure 11: $e_{c,total}$ for different combination of the received signals and databases with five algorithms: CC-T (Cross-correlation in the time domain), CC-F (Cross-correlation in the frequency domain), SCancellation (Source cancellation), SComparison (Source comparison), random



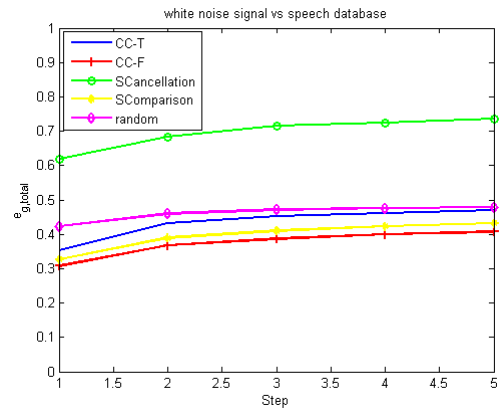
(a) white noise signals & white noise database



(b) speech signals & speech database

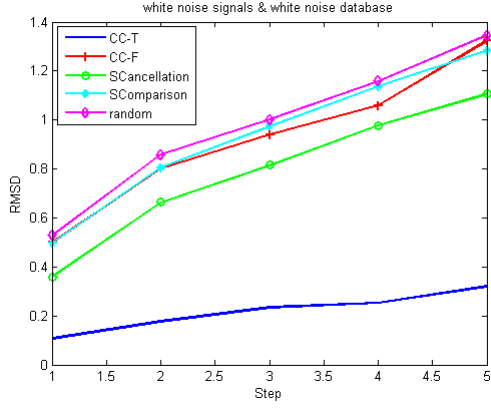


(c) speech signals & white noise database

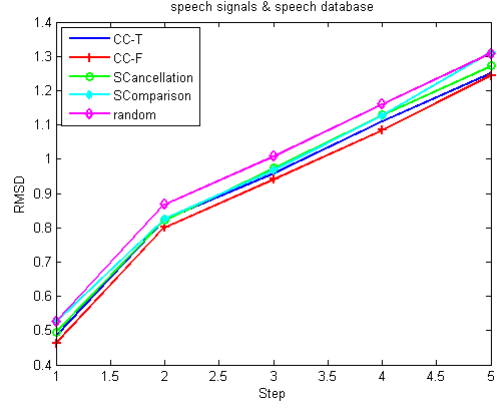


(d) white noise signals & speech database

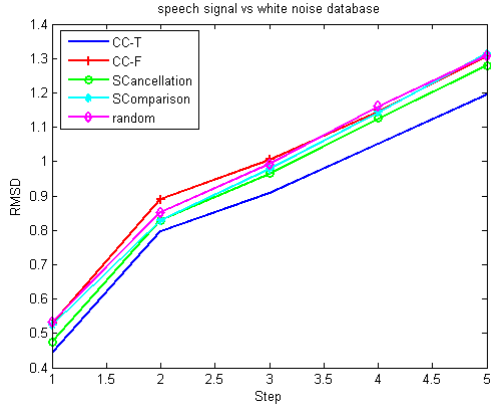
Figure 12: $e_{g,total}$ for different combination of the received signals and databases with five algorithms: CC-T (Cross-correlation in the time domain), CC-F (Cross-correlation in the frequency domain), SCancellation (Source cancellation), SComparison (Source comparison), random



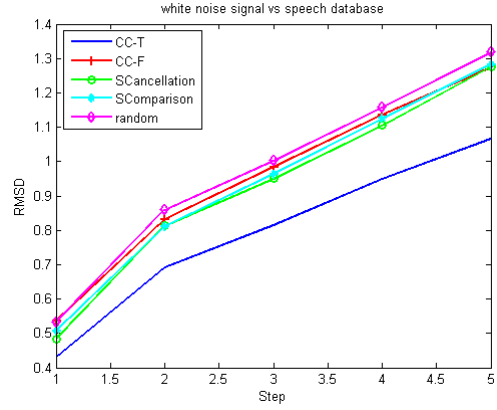
(a) white noise signals & white noise database



(b) speech signals & speech database



(c) speech signals & white noise database



(d) white noise signals & speech database

Figure 13: RMSDs in meter for different combination of the received signals and databases with five algorithms: CC-T (Cross-correlation in the time domain), CC-F (Cross-correlation in the frequency domain), SCancellation (Source cancellation), SComparison (Source comparison), random

The root mean square deviation of distance for one combination of types of signals is defined as follows,

$$\begin{aligned}
 RMSD_{d,total} &= \sqrt{\frac{\sum_{i=1}^{30} \sum_{j=1}^{56} d^2(i,j)}{30 \times 56}}, \text{ for the same type of signals,} \\
 RMSD_{d,total} &= \sqrt{\frac{\sum_{i=1}^{30} \sum_{j=1}^{64} d^2(i,j)}{30 \times 64}}, \text{ for different type of signals.}
 \end{aligned} \tag{52}$$

The results of RMSD are shown in Fig.13. This will be analysed in Sec. 7.

7 Discussion

In this part of the thesis all the tests will be reflected upon.

Test 1

Fig. 9 shows that after the back shifting, two signals can be matched. It illustrates that the shift can be found using the proposed algorithm. However, during the test, large and random deviations were presented due to the synchronization problem. The synchronization between two mobile phones was not well implemented at that time. Therefore among all the errors, the error introduced by bad synchronization was dominant. To reduce this error, a better synchronization is desired.

Test 2

The objective of the second test is to examine if the algorithm works with a reasonably good synchronization. From Tab. 1, it can be seen that with manual synchronization, the algorithm gives relatively small errors. In the practice, manual synchronization is not realistic. This test proves that the algorithm will give sufficient results given a reasonably good synchronization. The synchronization problem is usually expected to be solved via hardware/software implementation. However, if it can not be done, there is also possibility to estimate and compensate the errors from bad synchronization using statistic models. This approach is beyond the scope of this thesis and is thus not to be discussed in detail.

Test 3

The results are being revealed in Sec. 6. First of all each one of the four proposed comparison techniques is being analysed.

Cross-correlation in the time domain

When comparing Fig. 11(a), 11(b), 11(c) and 11(d), one is confronted with the following : the test results match the analysis stated in Sec.5 which implied that the cross - correlation in time domain would display the best performance in case of a white noise - white noise combination. The worst results concerning this particular method are observed in case of a speech - speech combination. The reason for this lies within the fact that white noise is uncorrelated with a particular shift in time while the speech signal is known for short term correlation. Due to this the side-lobes of the autocorrelation of the speech signal are relatively large when compared to those of white noise. The same conclusion can be drawn from Fig. 13(a), 13(b), 13(c) and 13(d).

Cross-correlation in the frequency domain

When comparing the results of the cross-correlation algorithm in the frequency domain as displayed in Fig. 11(a), 11(b), 11(c) and 11(d), it is relatively easy to note that the difference between the maximum correlation coefficient and the correlation coefficient at the desired point is rather small. However, the differences between various combinations of signals as conveyed by means of Fig. 11(a), 11(b), 11(c) and 11(d) do not correspond to Fig. 13(a), 13(b), 13(c) and 13(d). This originates from the standard deviation of the source signals and therefore there is no significant difference amongst the correlation coefficients. This conclusion is confirmed by the information represented in Fig. 12(a), 12(b), 12(c) and 12(d).

Source cancellation

From Fig. 11(a), 11(b), 11(c) and 11(d), a substantial increase of $e_{c,total}$ is observed in case of an increase of the given region. The relative error of the correlation coefficient is relatively large, except for the combination of white noise from the source and white noise from the database. Fig. 12(a), 12(b), 12(c) and 12(d) illustrate that the difference among the correlation coefficients is also quite large. Due to this stable results may be witnessed in 13(b), 13(c) and 13(d) while even better results are being displayed in 13(a).

Source comparison

From Fig. 11(a), 11(b), 11(c), 11(d), Fig.12(a), 12(b), 12(c) and 12(d) it is observed that the source comparison algorithm is quite stable amongst the four implemented combinations . The same conclusion can also be drawn from Fig. 13(a), 13(b), 13(c) and 13(d).

Finally, a comparison is being made between the four proposed techniques to find a match between an audio signal of the database and a random audio sample. When having a look at Fig. 13(a), 13(c) and 13(d) it becomes clear that cross-correlation in the time domain is the best algorithm amongst the four propositions. Its performance is not too bad as well in comparison to the other three methods according to Fig 13(b).

8 Conclusion

Indoor localization using signals of opportunity is an interesting and promising topic. One of the choices made during this work was to focus on acoustic signals which is caused by the fact that this particular type of signals has a much lower speed than the ones of an electromagnetic origin. Naturally, this implies that any possible timing inaccuracies would lead to smaller deviations in the localization process. This work placed the emphasis on a rectangular room as shown in Fig. 1 without loss of generality. The contents of the room are the following: a source producing an acoustic signal, a reference receiver and a mobile receiver. The main goal is the derivation of the position of the mobile receiver. A Wide-Sense-Stationary Uncorrelated Scattering model [13] is used to describe the room and the channel model is defined in (2) and (3). Back to the requirements described in Sec. 2, the requirements for the algorithm are sufficiently satisfied. The proposed algorithm is able to localize a mobile phone in a convex room using acoustic signals of opportunity.

The indoor localization algorithm proposed in this particular work is a combination of the Time Difference of Arrival and Fingerprinting. The first phase is the computation of the TDOA between the reference receiver and the mobile receiver which provides a possible region where the mobile receiver could be situated. The second phase consists of building database, signal processing and the comparison between the received signal from the mobile receiver and the one from the database. Four different approaches for signal processing have been presented.

Both phases of the suggested algorithm were thoroughly tested. From those particular tests it appears that the TDOA part is functioning properly but strongly depends on the synchronization. However it is demonstrated that the algorithm itself works sufficiently. Unfortunately, the synchronization issue was not really solved yet and was merely minimized by means of manual synchronization. Another possibility to diminish the timing problems was to come up with a stochastic model of the situation (e.g. estimation or detection theory), but our group considered such an analysis to be outside of the scope of the thesis especially due to the time limitations.

The tests concerning the Fingerprinting part of the proposed algorithm were mainly undertaken to decide which one of the four comparison methods would come out on top. Cross - correlation in the time domain has shown the best performance during this testing. It is also worthwhile to note that this particular part of the algorithm displayed the finest results in case of a white noise being used to construct the database. As already mentioned in Sec.4 and 5, this suggested way of using the Fingerprinting technique is relatively new which implies that a lot of testing needs to take place in order to be able to come up with reasonable assumptions concerning the performance of the proposed algorithm.

Although the proposed technique presented in this particular work requires additional testing the following conclusions may be drawn: the TDOA part of the algorithm does not improve on the state of the art while the Fingerprinting part introduces an interesting novelty with reasonable accuracy. The two parts working together ensures indoor localization with small deviations in an unprecedented manner.

9 Future work

In scientific research there is always room for improvement and this work is no exception. This particular extract proposes an algorithm for which there are elements that could be explored in more detail in the future.

The first element is the synchronization issue: although perfect synchronization does not exist, a sufficient synchronization can be achieved. This could be solved at the hardware level or at least some software diversions should be implemented. In case the above mentioned method is impossible to apply, a stochastic analysis of the situation should be performed.

The second element is the TDOA. Although the synchronization definitely plays a huge role during the implementation of this particular technique, the TDOA itself requires a thorough exploration as well. This work focussed on the usage of the white noise as well as speech signals, but there are many other possibilities. By having a look at other kinds of signals the robustness of the method might be improved. Another interesting topic concerning the TDOA which deserves future attention is the presence of multiple sources. This work focusses on one source only and it would therefore be worthwhile to consider the implications of the increment of the amount of sources.

The final element is the Fingerprinting. In this case two following aspects should be experimented with in the future: the kind of signal used to build the database as well as the amount of sources present in a particular scenario. The conclusion of this work implies that the Fingerprinting is performing at its best when a white noise signal is being used, but this might be proven wrong empirically. Also, as the main focus of this thesis was on a situation with a single source producing a signal, the implications introduced by multiple sources is the next logical step. And finally the last weak link of the Fingerprinting part: the applicability within the real world. It is not too hard to imagine the fact that not too many individuals would be pleased to be obliged to construct a database of each and every room where indoor localization would be applied later. A possible solution to this convenience problem would be the creation of a so-called virtual database which would make the long lasting measurements redundant. The idea would be to construct the database by means of simulation, taking the dimensions as well as any present object of the room into the consideration.

References

- [1] A. Tayebi, J. Gomez, F. Saez de Adana and O. Gutierrez, "The Application of Ray-Tracing to Mobile Localization Using the Direction of Arrival and Received Signal Strength in Multipath Indoor Enviroments" Progress In Electromagnetics Research, PIER 91, 1-15, 2009
- [2] G. Fischer, B. Dietrich and F. Winkler, "Bluetooth Indoor Localization System"
- [3] C. Zang, M. Kuhn, B. Merkl, A.E. Fathy and M.Mahfouz, "Accurate UWB Indoor Localization System Utilizing Time Difference of Arrival Approach"
- [4] Stefan Koenig, Mark T. Schmidt, Christian Hoene, "Multipath Mitigation for Indoor Localization Based on IEEE 802.11 Time - of - Flight Measurements"
- [5] A.H. Sayed, A. Tarighat, and N. Khajehnouri, "Network-based wireless location", IEEE Signal Processing Magazine, vol. 22, pp. 24-40, July 2005
- [6] E. D. Kaplan, C. J. Hegarty, "Understanding GPS: Principles and Applications", London, UK. : Artech House, 2006.
- [7] Fredrick S. Solheim, Jothiram Vivekanandan, Randolph H. Ware, Christian Rocken, "Propogation delays induced in GPS signals by dry air, water vapour, hydrometeors, and other particulates"
- [8] R.K. Martin, J.S. Velotta, and J.F. Raquet, "Bandwidth efficient cooperative TDOA computation for multicarrier signals of opportunity", IEEE Trans. Signal Processing, vol. 57, pp. 2311-2322, June 2009.
- [9] Vahideh Moghtadaiee, Andrew G. Dempster, Samsung Lim," Indoor localization using FM radio signals: a fingerprinting approach"
- [10] Lionel M.Ni, Dian Zhang, Michael R. Souryal," RFID - based localization and tracking technologies"
- [11] Oriol Vinyals, Eladio Martin, Gerald Friedland," Multimodal indoor localization: an audio - wireless - based approach "
- [12] Steadman, Philip. "Why are most buildings rectangular?." Architectural Research Quarterly 10.02 (2006): 119-130.
- [13] Kurth, R. R.; Snyder, D. L.; Hoversten, E. V. (1969) "Detection and Estimation Theory", Massachusetts Institute of Technology, Research Laboratory of Electronics, Quarterly Progress Report, No. 93 (IX), 177-205
- [14] Nicoli, M., C. Morelli, and V. Rampa. "A jump Markov particle filter for localization of moving terminals in multipath indoor scenarios." Signal Processing, IEEE Transactions on 56.8 (2008): 3801-3809.
- [15] Yano, Scott M. "Investigating the ultra-wideband indoor wireless channel." Vehicular Technology Conference, 2002. VTC Spring 2002. IEEE 55th. Vol. 3. IEEE, 2002.
- [16] Huang, Y-D., and Mourad Barkat. "Near-field multiple source localization by passive sensor array." Antennas and Propagation, IEEE Transactions on 39.7 (1991): 968-975.

- [17] D. Humprey and M. Hedley, "Super-Resolution Time of Arrival for Indoor Localization", 2008
- [18] Wang, X., Z. X. Wang, and B. O. Dea, "A TOA-based location algorithm reducing the errors due to Non-Line-of-Sight (NLOS) propagation," *IEEE Trans. Veh. Tech.*, Vol. 52, 112-116, Jan. 2003.
- [19] G. Gu, "Approximate MLE algorithm for source localization based on TDOA measurements", in *Proc. SPIE Conf.: Defense and Security 2007 (Orlando, FL)*, Vol. 6577, April 9 - April 13, 2007.
- [20] K.C. Ho and W. Xu, "An accurate algebraic solution for moving source location using TDOA and FDOA measurements", *IEEE Trans. Signal Processing* vol 52 pp 2453-2463. Sept. 2004.
- [21] Y.T. Chan and K.C. Ho, "A simple and efficient estimator for hyperbolic location", *IEEE Trans. Signal Processing*, vol. 42, pp. 1905-1915, Aug. 1994.
- [22] B.T. Fang, "Simple solutions for hyperbolic and related position fixes", *IEEE Trans. Aerosp. Electron. Syst.*, vol. 26, pp. 748-753, Sept. 1990.
- [23] A. Toak, N. Kandil and S. Affes, "Neural Networks for Fingerprinting-Based Indoor Localization Using Ultra-Wideband" *Journal Of Communications*, Vol. 4, No. 4, MAY 2009
- [24] Evgeny Kupershtein, Mati Wax, Israel Cohen, "Single - site emitter localization via multipath fingerprinting", *IEEE Transactions on signal processing*, Vol. 61, No. 1, January 1, 2013.
- [25] Self, Douglas (2012). *Audio Engineering Explained*. Taylor & Francis US. pp. 200, 446. ISBN 0240812735.
- [26] Butterworth, Stephen. "On the theory of filter amplifiers." *Wireless Engineer* 7 (1930): 536-541.
- [27] Li, Ta-Hsin, and Jerry D. Gibson. "Time-correlation analysis of nonstationary signals with application to speech processing." *Time-Frequency and Time-Scale Analysis, 1996.*, *Proceedings of the IEEE-SP International Symposium on.* IEEE, 1996.
- [28] A. Masoud, K. Warmerdam, "Implementation of Indoor Localization Using Signals of Opportunity on Android", TUDelft, bachelor thesis, 2013

A TDOA

extractTDOA.m

```
1 function[TDOA_meter]= extractTDOA(Signal_m, Signal_ref, max_meter, Fs)
2 % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
3 % Extract TDOA from two signals. %
4 % Signal_m: received signal from the mobile receiver. %
5 % Signal_ref: received signal from the reference receiver. %
6 % max_meter: the maximum possible difference. %
7 % Fs: sampling rate. %
8 % TDOA_meter: TDOA converted to length, which is the "distance difference %
9 % of arrival". %
10 % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
11
12 % speed of sound
13 c = 343.2;
14
15 % reduce the computation by setting up the maximum
16 Length_signal = min(length(Signal_m), length(Signal_ref));
17 percent = max_meter/c*Fs/Length_signal;
18
19 % compute TDOA
20 TDOA = comparevectorindex(Signal_m, Signal_ref, percent);
21 TDOA_meter = (TDOA/Fs)*c;
22
23 end
```

comparevectorindex.m

```
1 function [shift] = comparevectorindex(Sig1, Sig2, percent)
2 % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
3 % Compare two signals to find the shift. %
4 % Sig1: the first signal. %
5 % Sig2: the second signal. %
6 % percent: the percentage of the signal length which can be shifted. %
7 % shift: the sample shift found by this function. %
8 % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
9
10 % cut the signals to the same length
11 difference = length(Sig1) - length(Sig2);
12 if difference > 0
13     Sig1 = Sig1(1:end-difference);
14 elseif difference < 0
15     Sig2 = Sig2(1:end+difference);
16 end
17
18 % correlation
19 R = zeros(1, ceil(length(Sig2)*percent*2));
20 for i = ceil(-length(Sig2)*percent) : ceil(length(Sig2)*percent)
21     % shift one of signals
22     if i >= 0
23         X = [zeros(1,i) Sig2(1:end-i)];
24     else
25         j = -i;
26         X = [Sig2(j+1:end) zeros(1,j)];
27     end
```

```
28     % compute correlation coefficient
29     cof = corrcoef(Sig1,X);
30     R(i+length(Sig2)+1)=cof(1,2);
31 end
32 % find the best match
33 corcof_index = find(R == max(R));
34 shift = corcof_index-length(Sig2)-1;
35
36 end
```


B Fingerprinting

builddatabase_room_4.m

```
1 function [hr_left,hr_right] = Builddatabase_room_4(Type_signal)
2 % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
3 % Build the database. %
4 % Type_signal: white noise, speech... %
5 % hr_left: signal received by first channel of the receiver. %
6 % hr_right: signal received by second channel of the receiver. %
7 % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
8
9 % set up the size
10 Size_database = [2000, 3, 10];
11 hr_left = zeros(Size_database);
12 hr_right = zeros(Size_database);
13
14 % store the signals in the database
15 for x=0:2
16     for y=0:9
17         % read the .wav files
18         str_x=num2str(x);
19         str_y=num2str(y);
20         str_naam=strcat(str_x, str_y, '_', Type_signal);
21         temp_vec_2_bij_1=(wavread(str_naam))';
22
23         % store the signals
24         hr_left(:,x+1,y+1)=temp_vec_2_bij_1(1,:);
25         hr_right(:,x+1,y+1)=temp_vec_2_bij_1(2,:);
26     end
27 end
28
29 end
```

Compare_CC_T.m

```
1 function [corcof]= Compare_CC_T(hr_left, hr_right, Sig_left, Sig_right)
2 % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
3 % Compare the received signal with the database using Cross correlation %
4 % in time domain method. %
5 % hr_left: the database of the first channel. %
6 % hr_right: the database of the second channel. %
7 % Sig_left: received signal at the first channel. %
8 % Sig_right: received signal at the second channel. %
9 % corcof: a matrix contains correlation coefficients. %
10 % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
11
12 % set up the size
13 dims = size(hr_left);
14 corcof = zeros(dims(2), dims(3));
15
16 % cross correlation
17 X12 = xcorr(Sig_left, Sig_right);
18
19 % compare the received signal with the database
20 for i=1:dims(2)*dims(3)
21     X34 = xcorr(hr_left(:,i), hr_right(:,i));
```



```

11 % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
12
13 % set up the size
14 dims = size(hr_left);
15 X34 = zeros(dims);
16 corcof = zeros(dims(2), dims(3));
17
18 % received signals in the frequency domain
19 X1 = abs(fft(Sig_left));
20 X2 = abs(fft(Sig_right));
21
22 % divide the signals in the frequency domain
23 X12 = X1./X2;
24
25 % compare the received signal with the database
26 for i = 1:dims(2)*dims(3)
27     % the database in the frequency domain
28     X3 = abs(fft(hr_left(:,i)));
29     X4 = abs(fft(hr_right(:,i)));
30
31     % dividing
32     X34(:,i) = X3./X4;
33
34     % correlation coefficient
35     tmp = corrcoef(X12, X34(:,i));
36     corcof(i) = tmp(2);
37 end
38
39 end

```

Compare_SComparison.m

```

1 function [corcof] = Compare_SComparison(hr_left, hr_right, Sig_left, ...
2                                     Sig_right)
3 % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
4 % Compare the received signal with the database using Source comparison %
5 % method. %
6 % hr_left: the database of the first channel. %
7 % hr_right: the database of the second channel. %
8 % Sig_left: received signal at the first channel. %
9 % Sig_right: received signal at the second channel. %
10 % corcof: a matrix contains correlation coefficients. %
11 % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
12
13 % set up the size
14 dims = size(hr_left);
15 X3 = zeros(dims);
16 X4 = zeros(dims);
17 corcof = zeros(dims(2), dims(3));
18
19 % received signals in the frequency domain
20 X1 = abs(fft(Sig_left));
21 X2 = abs(fft(Sig_right));
22
23 % database in the frequency domain
24 X3 = abs(fft(hr_left));
25 X4 = abs(fft(hr_right));
26

```