On Aspects of Indoor Localization

Bei der Fakultät für Elektro- und Informationstechnik
der
Universität Dortmund
zur Begutachtung eingereichte

Dissertation

zur Erlangung des akademischen Grades
Doktor der Ingenieurwissenschaften

von

Holger Linde

August 2006

Zur Begutachtung eingereicht am 30.09.2005
Tag der mündlichen Prüfung: 07.06.2006

Hauptreferent: Prof. Dr.-Ing. U. Schwiegelshohn
Korreferent: Prof. Dr. M. ten Hompel

Abstract

One of the key issues of emerging mobile computing and robotics is to obtain knowledge of the position of persons, vehicles and objects in an indoor environment. In order to enable a pervasive coverage, this must be achieved at low cost. Due to the importance of this problem, numerous solutions have been proposed. In this thesis two important localization techniques are discussed and major improvements are developed to address their weaknesses.

A common approach for the position estimation problem is hyperbolic localization. This method is based on time difference of arrival (TDOA) measurements of signals transmitted from the mobile unit, i.e. the unit that is to be located, to a set of fixed reference units. A key requirement of this technique is that the clocks of the reference devices are synchronized. Unfortunately, if electromagnetic signals are used that travel at the speed of light even a very small clock deviation results in a tremendous inaccuracy of the computed location. Therefore, a highly precise time measurement and synchronization is mandatory for these systems. However, clocks operating at the required resolution and precision are complex and thus expensive. In this thesis, a localization approach based on TDOA measurements that implies clock synchronization is proposed and discussed. It allows for the usage of low cost oscillators that operate at a moderate frequency in the order of 100 MHz. The impact of short and long term instabilities like jitter or clock drift are inherently delimited. The approach can be applied to radio or infrared as well as ultrasound based systems. The main advantage of this approach is that common synchronization mechanisms that require a significant amount of processing and/or hardware resources can be neglected. Hence, this method is well suited for applications where the mobile unit must be very low cost and thus, of low complexity.

An alternative to hyperbolic localization is triangulation, which requires the measurement of angles among fixed reference units and the mobile target. In this work, an infrared detector array for angle of arrival measurement is presented. The array consists of multiple sensing elements that are orientated in different directions. First, the arrangement is described by a sampling system. However, it is shown that low cost integrated receivers yield various aberrations, and thus the sampling approach fails. Subsequently, a new approach named virtual filter interpolation is proposed and discussed. This approach can handle individual sensitivity characteristics of each sensor element and therefore outperforms the sampling approach. The proposed technique is qualified for extremely low cost localization, e.g. for robot navigation.

Acknowledgement

The work for this thesis has been carried out in the Section Information Technology at the Robotic Research Institute (IRF) in Dortmund, Germany. I wish to express my sincere gratitude to the people that supported me during the last couple of years. First of all I want to thank my supervisor Uwe Schwiegelshohn for giving me the opportunity to work on this project, and for his invaluable advice and encouragement. Furthermore I wish to thank Edwin Naroska, who was an inexhaustible source of inspiration and helped whenever I was in need. Without our countless discussions I had certainly been stuck in blind alleys several times. His presence at the institute was the best thing that could happen to me and my work.

During this work, I have received a three years scholarship from the Graduate School of Production Engineering and Logistics (GSPEL), an institution of the University Dortmund. For this funding, I am very thankful. I would also like to take this opportunity to thank my colleagues for creating an ever nice and friendly atmosphere in the institute. I enjoyed working with them very much. Lastly, I am very grateful for the love, support and patience of my parents and my dear Johanna.

Contents

List of Figures

List of Tables

Definitions

This section summarizes some denotations and definitions used throughout the following chapters.

- Functions, matrices, vectors and angles are denoted as follows:

Type	*Examples*
Functions	$f(x)$
Vector functions	$\vec{f}(x)$
Derivatives	$f'(x)$, $\vec{f}'(x)$
Column vectors	$\mathbf{r}$, $\mathbf{R}$
Row vectors	$\mathbf{r}^{\mathrm{T}}$, $\mathbf{R}^{\mathrm{T}}$
Vector elements	$r(n)$, $R(k)$
Matrices	$\underline{\mathbf{A}}$
Matrix elements	$a(i,j)$, $a_i(j)$
Angles	α

- Discrete signals are indicated by a lower case character (e.g., $\mathbf{r}$, $r(n)$) when considered in the angular domain and an upper case character (e.g., $\mathbf{R}$, $R(k)$) for the frequency domain.

- The measure of angles is degree (°) if not mentioned otherwise.

- The frequency spectrum $S(f)$ of a continuous signal $s(\phi)$ is defined by the forward and inverse Fourier transforms, denoted FT and FT^{-1}, respectively:

$$S(f) = \mathsf{FT}\left[s(\phi)\right] = \int_{-\infty}^{\infty} s(\phi) \cdot e^{-j2\pi f\phi}\, d\phi \tag{1}$$

$$\Updownarrow$$

$$s(\phi) = \mathsf{FT}^{-1}\left[S(f)\right] = \int_{-\infty}^{\infty} S(f) \cdot e^{j2\pi f\phi}\, df \tag{2}$$

Note that in Eq. 1 and 2, ϕ is measured in radian (rad). Other than that, throughout this work, angles are measured in degrees (°).

- The frequency spectrum of a signal defined in the *angular domain*, $s(\phi)$, is denoted $S(f)$. f is given in $1/1°$.
- Accordingly, the forward and inverse discrete Fourier transforms (DFT and IDFT), here denoted DFT and DFT^{-1}, respectively, are defined as:

$$S(k) = \mathsf{DFT}\left[s(n)\right] = \sum_{n=0}^{N-1} s(n) \cdot e^{-j\frac{2\pi}{N}\cdot k\,n}, \quad k \in \{0, 1, \ldots, N-1\} \tag{3}$$

$$\Updownarrow$$

$$s(n) = \mathsf{DFT}^{-1}\left[S(k)\right] = \frac{1}{N}\sum_{k=0}^{N-1} S(k) \cdot e^{j\frac{2\pi}{N}\cdot k\,n}, \quad n \in \{0, 1, \ldots, N-1\} \tag{4}$$

- The rectangular function rect(x) is defined as

$$\operatorname{rect}\left(\frac{x}{X}\right) = \begin{cases} 1, & \text{if } -\frac{X}{2} < x < \frac{X}{2} \\ \frac{1}{2}, & \text{if } |x| = \frac{X}{2} \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

- The operation $\mathbf{a} \div \mathbf{b}$ indicates element-wise division of two vectors $\mathbf{a}$ and $\mathbf{b}$.
- The operation $\mathbf{a} \otimes \mathbf{b}$ indicates cyclic (circular) convolution of the sequences $\mathbf{a}^{\mathrm{T}}$ and $\mathbf{b}^{\mathrm{T}}$.

Chapter 1

Introduction

Ever since mankind started to explore its habitat, the question "Where am I?" has arisen. No matter if it was about finding again recently discovered water or food resources, or simply the way back home: Estimating one's own position has always been a central problem in human history. The peoples of the earth developed different strategies to answer the above question: They analyzed the environment and memorized striking features, such as river shores, mountains, or specially shaped trees. When moving in an explored area, they obtained their position by recognizing these features. Also analyzing the constellation of celestial bodies was early discovered to be a reliable tool. Besides observing their surroundings and the sky, humans have always orientated themselves by inherently keeping record of their recent movements, in a fashion of "a short distance straight ahead, then a few steps to the right". The process of navigation for the human mind has always been an iterative one. At first, people have a coarse impression of their whereabouts. Then, by recognizing more and more features of the environment, the location estimation is improved.

During the evolution of mankind, much scientific and technological progress has been made: The achievements in mathematics, headed by the laws of trigonometry, and the invention of novel measuring instruments have paved the way for new localization procedures with constantly growing accuracy. These techniques were used in old and emerging applications, including terrestrial and naval navigation as well as geodesy. Aircraft and space craft navigation, guided missiles and robot navigation are a few of the more recent applications depending on localization.

While the need for estimating the physical position of humans and objects has accompanied mankind since its first steps, it becomes apparent that in the future, the importance of localization will even grow. New breakthroughs, particularly in micro electronics, sensors and information technology, enable ever new applications that demand location information at an ever growing accuracy or lower cost. Furthermore, many of these applications are situated indoors.

The range of traditional and emerging applications requiring indoor position data is man-

ifold. A large classic part is in the industrial and logistics sector. *Automated* or *self guided vehicles* (AGV/SGVs) and *robots* are instructed to transport goods in an automated production facility or warehouse. In order to navigate autonomously or by remote control, they must localize themselves in their surroundings. Besides these traditional applications, indoor localization has experienced additional efforts in recent years with the appearance of ubiquitous computing. In various scenarios *objects* and *goods* must be located or tracked, e.g. in an industrial or medical environment. Additionally, *people* localization enables a large number of applications and services. Customers and employees can be observed, intruders detected, elderly people supported, and patients tracked. In public places like subway stations, safety concerns can be addressed by an emergency guidance system. A similar situation may occur in a hotel. A guidance system fed with position information of humans can work more efficiently than a static one. Another application that has raised much interest in recent years is ad hoc sensor networks. Besides distributed sensing it allows for location aware routing which may improve network efficiency and lower power demands. A few applications for indoor localization are summarized in Table 1.1.

Industrial	Logistics	Commercial	Medical	Smart house	Public places	Tourism	Sensor networks	Target	Purpose
☒	☒	☐	☐	☐	☐	☐	☐	AGVs and robots	Navigation in production facility or warehouse
☒	☒	☒	☒	☒	☐	☐	☐	Goods	Finding goods/objects
☐	☐	☒	☐	☐	☐	☐	☐	Customers	Profiling customers' habits, shopping assistance by navigation
☐	☐	☐	☐	☒	☐	☐	☐	People	Smart house climate control, convenience features, intruder detection, electronic home arrest monitoring
☐	☐	☐	☒	☐	☐	☐	☐	Patients	Finding patients
☒	☐	☒	☒	☐	☐	☐	☐	Employees	Observation and profiling
☐	☐	☐	☐	☐	☒	☒	☐	People	Emergency guidance systems
☐	☐	☐	☐	☐	☐	☐	☒	Nodes	Location aware routing, distributed sensing

Table 1.1: Examples of applications utilizing indoor localization

By nature, positioning is an interdisciplinary problem that puts forth countless questions in various research fields, such as engineering, computer science and statistics. As a consequence, designing and implementing a localization system is a very complex task implying an understanding in these fields. However, the methodology of location estimation is in principle the same that was used in ancient times. Where people used the constellation of stars to deduce location, today we deploy fixed reference devices at known locations of our surroundings. Then, the mobile target whose position is to be determined measures angles or

distances towards these reference points. From these measurements, the absolute position can be determined by triangulation techniques. Likewise, in the same way in which humans keep record of their recent movements, the position of the mobile target can be determined by course and velocity information obtained from rotary encoders or accelerometers. The same techniques and algorithms can be applied for both outdoor and indoor localization. However, the typical indoor environment contains multiple walls and floors and a large number of objects (e.g., furniture) that are made of different materials. This cluttered environment evokes fundamentally new challenges. Besides the environment, most indoor applications have significantly higher demands concerning localization accuracy, cost and update rate than most outdoor applications. As a result, hardly any existing outdoor localization solution operates to a satisfying degree in the cluttered indoor environment. Hence, existing approaches must be improved, or new designs developed, especially under the condition to lower cost.

Depending on the localization approach, certain fundamental problems exist that must be solved in every implementation. These problems are discussed in detail in the following chapter. For now, we can state that they arise from the fact that measurements are always afflicted by errors. There are three options to improve the quality of measurements: either use complex and thus costly measuring methods, design workarounds, or develop new measuring approaches. In this thesis, different methods are proposed to achieve a *high measurement accuracy* at *low cost*. Therefore, all methods presented here can be implemented using inexpensive commercial off-the-shelf (COTS) components.

The work is structured in the following way: In Ch. 2, definitions and performance measures of localization systems are introduced. Based on the indoor applications mentioned above, common properties and demands on the localization system are given. Based on the classification of localization approaches, common localization techniques are introduced: *trilateration*, *hyperbolic localization* and *triangulation*. Subsequently, common signal technologies like ultrasound, infrared and radio are discussed. Advantages, drawbacks and challenges of the covered techniques and technologies are discussed. In Ch. 3, state of the art implementations are briefly discussed and evaluated. Ch. 4 deals with a new method for high accuracy synchronization and time measurement in hyperbolic localization systems. In Ch. 5, a direction finding algorithm based on a low cost infrared detector array is presented. The work is concluded in Ch. 6 with a summary and outlook.

Chapter 2

Basics of localization

2.1 Common properties of indoor localization systems

The applications addressed in this work have similar, yet not identical demands on the localization system. Prior to choosing an appropriate versatile positioning technology, common properties must be analyzed. In the subsequent step, the required performance according to the individual application must be analyzed.

- The environment where localization is performed is known beforehand and can therefore be prepared to actively support positioning. That is, fixed reference units can be deployed in this environment, from the interaction of which the location of a mobile target can be derived.
- A 2D or 3D geometric location model can be applied to describe location information. This would typically be a local Cartesian or polar coordinate system, delivering x, y and z coordinates, or latitude, longitude and altitude, respectively. In certain applications, a context based, logical location model might be more appropriate. E.g. to express that a person is situated in a certain room, one can use a label (*kitchen*, *living room*, etc.) to refer to this room. This logical model can be derived from the geometric location model.
- The system's *scale* does not exceed a few hundred meters. Localization systems limited to a relatively small area are often referred to as *local positioning systems* (LPS), in order to explicitly distinguish them from large scale implementations like 2G/3G cellular network positioning systems, or global navigation satellite systems (GNSS). Indoor systems fall into the LPS category by nature.
- In all cases, positioning can be performed using relative or absolute localization methods, or both[1].

[1]The terms *relative* and *absolute* positioning are discussed in Sec. 2.3.

- Localization is performed in air.
- There may be differences regarding the environmental[2] and ambient[3] conditions where positioning must be performed, not only between the above mentioned scenarios but also locally inside one and the same application. These changes must be taken into account when choosing the localization technology.

2.2 Demands and performance measures in indoor localization

The applications introduced above make similar, yet not identical demands on the localization system. In order to judge an implementation, its performance parameters must be analyzed. These measures are listed in the following.

- One of the major performance parameters of a localization system is its *accuracy*. This term indicates how far the estimated location deviates from the true location. Accuracy is often confused with *precision*. Precision is a measure of the reproducibility of a measurement. Therefore, when a statement is made about the resolution of a localization system, accuracy is commonly supplemented by precision (e.g. 20 cm accuracy over 90% of the time). For the applications mentioned above, the required localization accuracy ranges roughly from ±5 cm (fine-grained robot positioning) to about ±1 m (coarse-grained guidance system). In some cases, even less accuracy is sufficient, for instance in occupancy detection in a room.
- *Cost* is a critical issue, particularly in applications where numerous units have to be equipped with localization technology, such as ad hoc sensor networks. It is therefore reasonable to use inexpensive standard off-the-shelf components whenever possible. Besides the mobile targets, infrastructure, deployment and maintenance cost must be considered.
- *Power consumption* must be considered, especially when a large number of battery driven mobile units are involved.
- Another point is *scalability*. On the one hand, the system's outer borders must be easily extendable (e.g. by adding further rooms to the covered area). On the other hand, the system must be able to cope with an increasing number of mobile targets that need be located concurrently.
- The *estimation cycle length*[4], i.e., the time between two location updates, must be considered in some cases. For instance, autonomous robots will act blindly[5] until the

[2] e.g. different walls, resulting in a varying grade of reflection, scattering and attenuation
[3] e.g. temperature and humidity
[4] also referred to as *responsiveness*
[5] unless they are equipped with odometric sensors

next position update. If the update cycle is too long, they can only operate reliably at a very low speed in order to avoid collisions.

- *Robustness* is another aspect that need be mentioned. This includes the level of immunity against interference from other sources, but also from ambient variations such as temperature and humidity etc.

- Particularly in those systems where active tags are carried around by humans, *device size* must be taken into account. If these are large and heavy, the positioning system might be turned down by its users for concerns of convenience.

- The utilized technology must be *compatible* with human beings that work or live within range of the localization system. For instance, acoustic range measurement using artificial sound sources is unacceptable. Many domestic animals can perceive sound at frequencies much higher than humans. Hence they might be affected by systems using ultrasound.

In the past decades, a number of different positioning approaches have been developed and published. As by now, there is not "the one" superior concept that performs best in each and every situation and application. Instead, most methods have been developed and enhanced for one particular purpose, so while, for example, one system may work perfectly in an outdoor environment it may be disturbed or malfunction indoors. For example, existing large scale infrastructures like cellular phone systems or global navigation satellite systems like GPS usually lack the required accuracy [18, 24, 45] for indoor applications. Furthermore, the GPS signal is attenuated by walls to a degree that it is hardly available indoors. In recent years, numerous dedicated indoor localization systems have been presented. Some of these systems are briefly discussed in Sec. 3. Prior to evaluating these real world implementations, a survey of general positioning methods is given.

2.3 Relative and absolute positioning

Since there exist many potential ways to obtain location information, it is reasonable to classify localization techniques. One such classification is based on the type of sensors utilized, and thus the way of processing the gathered sensor data in order to compute the location of the target. Roughly, this approach puts forth two classes of positioning systems, relying on either relative or absolute data acquisition. In Sec. 2.3, a number of general concepts will be introduced and briefly discussed. A more detailed survey can be found in the work by Borenstein et al. [13]. There exist a number of other classifications, one of them being the discrimination between indoor and outdoor positioning. For further reading on the taxonomy of localization systems the reader is referred to the work of Muthukrishnan et al. [66]. In this thesis, the classification of relative and absolute positioning is used.

2.3.1 Relative positioning

Relative positioning, also known as *dead reckoning*[6], describes the procedure of determining the current location of a mobile target by using course and velocity information. It can be subdivided into two approaches:

Odometry

Odometry is a very straight forward approach for dead reckoning. Angular position data collected from rotary encoders placed at motor axles or wheels is recorded for a given length of time. Starting from a known position, the present location of a unit can then be determined by reconstructing the covered path. Odometry is fully self-contained but on the other side susceptible for errors caused by e.g. wheel slippage.

Inertial Navigation

Inertial navigation systems (INS) use gyroscopes and accelerometers to measure the rate of rotation and acceleration of a unit. Position information is computed by integrating the measured data twice. Like odometric systems, inertial navigation is self-contained, but in the same way errors can accumulate without bounds. An inertial system can help to compensate momentary odometric errors.

The joint method utilizing both odometric and inertial sensors is sometimes referred to as gyrodometry [14].

2.3.2 Absolute positioning

The absolute position of a mobile unit can be determined with the aid of fixed reference points located in the environment. The position of these reference points is known a priori. These references can be active or passive components. A good survey on this topic covering historical and technical aspects as well as some implementations can be found in the work of Jiménez et al. [50]. This section introduces three approaches: *Beacon registration*, *landmark recognition* and *model matching*.

[6] derived from "deduced reckoning", used in naval navigation.

Beacon Registration

Active beacons[7] are static components located at fixed and known positions of the environment[8]. There are two different types of beacons: self-acting beacons that periodically emit a certain signature (e.g., a unique bit sequence), and sensitive beacons. Sensitive beacons either act as listeners or actively reflect a received signature emitted by the mobile unit. Beacon signals can be used for localization based on trilateration or triangulation techniques (see Sec. 2.4.1 and 2.4.6). Depending on the target application, ultrasound, RF or infrared transmission is used for range or angle estimation. The actual distance between a beacon and the mobile unit can be determined by either *time of flight* (TOF), *phase shift* or *signal attenuation* measurement. For example, TOF is computed in GPS receivers, while many laser ranging systems are based on phase shift measurement. Precise distance measurement using electromagnetic waves is tricky due to the high signal propagation speed($\approx 1\,\mathrm{cm}/33\,\mathrm{ps}$ in air). Implementations of this kind are discussed in Sec. 3.

Landmark Recognition

Landmarks are static features of an environment that can be recognized by a mobile unit. Most of the time, landmarks are geometric shapes like rectangles, lines or barcodes. Besides these artificial objects, also natural items such as doors can serve as landmarks. Also, the general term *scene analysis* is often used in this context. Mobile units trying to localize themselves in an environment use camera vision and feature extraction to analyze the current scene. Once a landmark has been reliably recognized and identified, the unit's current position relative to the corresponding landmark can be calculated. This is accomplished by triangulation and trilateration as well as other methods. For this landmark recognition, no *active* components must be deployed in the environment. On the other hand, landmarks must be carefully designed and placed in order to guarantee reliable operation. This includes, for example, sufficient contrast (and possibly hue) relative to the background. An insufficient design and setup may lead to problems as soon as ambient conditions such as lighting change. Also, considerably more processing resources are necessary for landmark based positioning than for active beacon systems. As all optical approaches, these systems suffer at situations where an obstacle is located between the camera and the landmark, thus shadowing the landmark.

Model matching

In some applications, a mobile unit must be able to build a map or world model of an unknown environment while at the same time localizing itself within that map. While moving and exploring, a reference world model is created. Positioning is accomplished by comparing this

[7] deduced from conventional beacons known from rail and ship traffic.

[8] Note that exceptions exist: Global localization systems like GPS are based on satellites orbitting the planet. In this case beacons are not static.

(possibly pre-stored) reference model to a local model generated from on-board sensor data. Thus, two issues must be addressed: *map building* and *map matching*. Map building can also be accomplished in an offline training phase. Features of the environment are sometimes called *fingerprints*. Typically, no single sensor technique can collect all relevant data in a natural environment to guarantee reliable operation. Heterogeneous sensor systems are hence combined, for example laser radar, a sonar and an odometric system. Received radio signal characteristics such as signal strength or multipath patterns are also commonly used. Model matching is rather costly regarding the demands on processing, memory and therefore energy resources.

2.3.3 Comparison

As mentioned before, all of the presented methods have certain strengths and weaknesses. The major drawback of relative positioning systems is their susceptibility for unbounded accumulated errors. So while easy to implement in most cases, dead reckoning should not be used as a stand alone solution. Instead, it may be combined with an absolute localization system. Regarding absolute position determining, both landmark recognition as well as model matching have very high demands on the processing performance of the mobile unit and are hence costly. Contrary to this, trilateration/triangulation based localization using active beacons can be performed with less computing efforts. In the following sections, these techniques will be discussed in more detail. This work focuses on aspects of absolute positioning. Therefore, relative localization methods are not covered.

2.4 Localization schemes

2.4.1 Trilateration

Trilateration is a technique to compute a position of an object m, given its distances l_{am}, l_{bm} and l_{cm} to three fixed non-collinear[9] reference objects a, b and c (see Fig. 2.1). For every distance l_{pm} between m and p (with $p \in \{a, b, c\}$), a circle at (x_p, y_p) with radius l_{pm} can be drawn around p. The point of intersection of three of these circles then yields the coordinates (x_m, y_m) of m. Therefore, trilateration can be expressed as finding the solution to the following system of quadratic equations:

$$\begin{aligned}(x_m - x_a)^2 + (y_m - y_a)^2 &= {l_{am}}^2 \\ (x_m - x_b)^2 + (y_m - y_b)^2 &= {l_{bm}}^2 \\ (x_m - x_c)^2 + (y_m - y_c)^2 &= {l_{cm}}^2\end{aligned}$$

[9] shown here in 2D. 3D localization requires four non-coplanar references

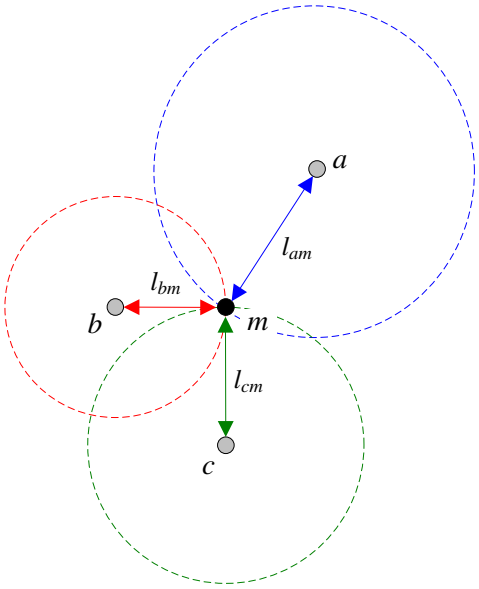


Figure 2.1: Trilateration scheme

A closed-form solution is discussed by e.g. Coope [23]. Under certain conditions, it may suffice to consider only two circles. In this case, an ambiguity exists because the system yields two solutions. If the approximate position of the mobile unit is known, the true position may be found by choosing the more plausible option. A solution to the trilateration problem in 3D is presented by Manolakis [60], which is considered as the computationally most efficient method for the trilateration problem. There also exist numerical resolution approaches.

Relying on the distances l_{pm}, trilateration requires the *measurement of distances* between the mobile unit and the reference units. This can be achieved by measuring either of the following:

- Time of flight (TOF)
- Phase shift
- Signal strength

These techniques are further described in Sec. 2.4.2, 2.4.3 and 2.4.4.

2.4.2 Time of arrival measurements

In order to obtain the range between two units m and p, beacon signals are transmitted between these devices. Typically, for time of flight range measurements a clock is needed on both units. If these clocks are synchronized, and if the time of emission of the beacon signal is

known, then from the time of arrival (TOA) the time of flight (TOF) t_{pm} can be determined. Multiplying by the signal propagation speed v yields the distance between the objects, l_{pm}.

In many applications one of the objects is not equipped with an appropriate clock at all, or the clocks are not synchronized. In this case, the distance l_{pm} between the two units can be determined by sending a beacon signal from m to p and measuring the roundtrip time $\Delta t_{pm} = 2\,t_{pm} + t_r$, given that the signal is reflected by object p:

$$l_{pm} = \frac{v \cdot (\Delta t_{pm} - t_r)}{2}$$

The beacon signal may be passively (reflection at surface) or actively (retransmission) reflected. The passive reflection technique is well known as RADAR (radio detection and ranging), which is used in large scale outdoor environments. When optical radiation or acoustic waves are used, the terms LIDAR and SONAR (light detection and ranging / sound navigation and ranging), respectively, are applied. Passive reflection indoors is not feasible for most signal technologies (e.g. RF) because of the high degree of multipath occurrences (i.e., echos due to signal reflection at walls and objects). For an active reflection, a beacon retransmission latency t_r must be considered. In the passive case $t_r = 0$ holds.

If active beacons are deployed in the area, commonly the one-way measurement approach is applied, i.e. the units must be synchronized. In case of radio signals, synchronization of the clocks is critical. Due to the high propagation speed, even a small synchronization deviation may result in huge measurement errors. This issue is further discussed in Sec. 2.7.

2.4.3 Phase shift

The time of flight between two objects p and q can also be determined using a continuous periodic signal, e.g. a (periodically) modulated RF carrier. The signal generated and transmitted by unit p is after the time of flight received by unit q. Internally, unit q generates the same signal and performs a cross correlation between the internal and the received signal. If the units are perfectly synchronized, i.e. the signals are generated concurrently, the result of this operation yields the phase difference ϕ of the two signals. This phase difference is proportional to the distance between the two objects, d. The distance can then be computed as

$$d \;=\; v \cdot T \cdot \frac{\phi}{2\pi},$$

where v is the signal propagation speed and T the signal period. To avoid any ambiguities, $d < v \cdot T$ must hold.

If the signal is passively reflected by an object, no synchronization is required, and the roundtrip time can be measured. This type of roundtrip time measurement is often used with mobile robots. These robots scan their environment for obstacles using rotating infrared laser beams or ultrasonic waves. The modulated emitted signal is autocorrelated with the reflected received signal. The time of flight corresponds to the phase shift of the modulation signal.

2.4.4 Received signal strength

A common ranging approach that goes without intricate clock synchronization is based on signal attenuation. That is, range is deduced from the received signal strength (RSS). However, indoor RF signal strength is non-linear with distance and covered with non-Gaussian noise as a result of multipath effects and environmental effects such as building geometry and traffic. A number of empirical models have been set up (e.g., [74]). However, these models strongly depend on signal properties and on the environment in which measurements are performed. Due to the large uncertainty of the path loss model RSS rarely provides the accuracy required for indoor range estimation. Instead of ranging, in most implementations exploiting RSS information, a map of *fingerprints* is built in a training phase. During operation, measurements are correlated with the map data in order to estimate location. The general limits of localization using signal strength are discussed by Elnahrawy et al. [26].

2.4.5 Hyperbolic localization

Instead of acquiring times of arrival (TOA) between reference units and the mobile unit, commonly time *differences* of arrival (TDOA)[10] are measured. The setup of the resulting localization principle which is commonly referred to as *hyperbolic localization* is shown in Fig. 2.2.

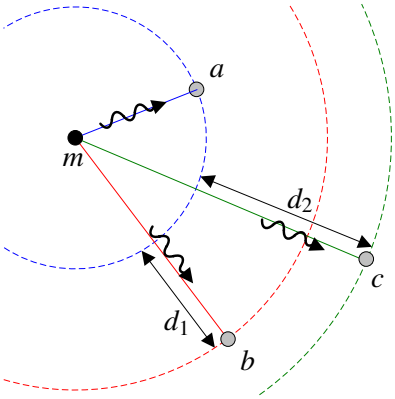


Figure 2.2: Hyperbolic localization scheme

Signals emitted from the mobile unit m are received by three reference units a, b and c. l_{pq} denotes the physical distance between nodes p and q, with $p, q \in \{a, b, c, m\}$). For the exact solution in the plane, two TDOA measurements are required which yield the distance differences d_1 and d_2. In 3D, three TDOA measurements (i.e., four reference units) are necessary. Unlike TOA measurements, TDOA measurements are independent of the signal emission time. Therefore, only the reference units must be synchronized among each other.

[10] sometimes also referred to as DTOA (Differential time of arrival)

The term *hyperbolic* arises from the fact that each distance difference yields a hyperbola (a hyperboloid if three dimensions are considered). The intersection of these hyperbolas yields the target position. A graphical analysis illustrating hyperbolic localization is given in App. A. It is difficult to find a closed-form solution of hyperbolic intersection, i.e., to find the position of the mobile target, (x_m, y_m), from the distance differences d_1 and d_2. Alternatively, the *direct calculation* method as shown in App. B can be performed. If additional information besides the TDOA measurements (e.g., altitude) is available, the complexity of the solution can be reduced [27]. When redundant reference units are deployed, there exists an overdetermined equation system. In this case, the direct method can be employed for different sets of reference units, thus generating multiple solutions. The final estimate may then be chosen by computing the average, or by selecting the set that yields the highest signal to noise ratio (SNR).

As an alternative to the direct calculation method, the system may be linearized by Taylor series expansion and solved iteratively by determining the least squares (LS) solution [34, 86]. Unfortunately, this approach depends on an initial location estimate close to the true position which in practice may not be available. Also, it is computationally intensive as LS computation is required in each iteration. Several alternative methods have been published to avoid this drawback. They can also cope with redundant reference units. In the work by Schau and Robinson [76], Friedlander [36] and Smith and Abel [81], the hyperbolic intersection problem is recast into a spherical intersection problem. Chan and Ho proposed maximum likelihood estimation [21].

A common drawback of the direct hyperbolic positioning approaches mentioned so far is that they yield multiple solutions. The Taylor series method does not – but on the other hand it requires an initial location estimate. An alternative localization paradigm exhibiting none of these constraints is using a non-linear optimization strategy. Several methods have been employed for iterative estimation, including the Gauss-Newton method, the Levenberg-Marquardt method and the quasi-Newton method. In context of the latter, a number of algorithms have been examined, e.g. the DFP (Davidon-Fletcher-Powell) minimization algorithm [29] and the BFGS (Broyden-Fletcher-Goldfarb-Shanno) algorithm [17]. The DFP algorithm has for instance been employed in the UWB localization system by Multispectral Solutions [32].

2.4.6 Triangulation

Localization based on trilateration implies that participants must be able to measure distances (or distance differences) between each other. Triangulation works in a similar manner, but instead of distances, angles are measured. In fact, it can be shown that triangulation can be transformed to trilateration by simple means. Two cases can be distinguished, as illustrated in Fig. 2.3.

- The mobile unit measures angles towards signals emitted by fixed reference units (see Fig. 2.3(a)). The collected data yield position and orientation.

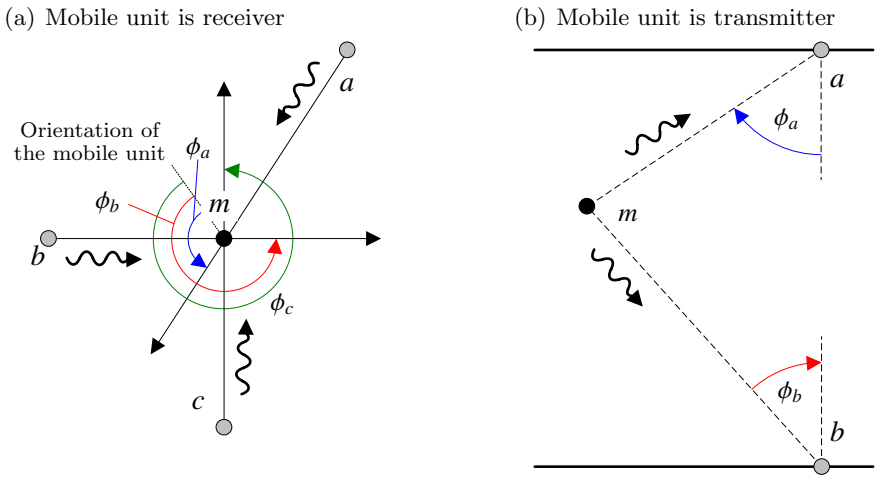


Figure 2.3: Triangulation schemes

- Reference units measure angles towards the signal emitted by the mobile unit (see Fig. 2.3(b)). Only a location estimate, but not the orientation of the mobile target, can be obtained. However, compared to hyperbolic localization, only two reference units need be deployed (in 2D), instead of three[11]. This arrangement is discussed in detail in Sec. 5.1.

In analogy to the terms TOA and TDOA, AOA denotes the angle of arrival. Several aspects concerning angle measurement are discussed in Sec. 2.4.7.

2.4.7 Angle of arrival measurement

The process of direction finding (DF) is usually accomplished by an *angle of arrival* (AOA)[12] measurement. The particular measurement arrangement depends on the signal technology used for localization. For instance, many sensing elements like ultrasonic transducers or infrared photodiodes have directional receiving patterns while e.g. rod antennas do not. Therefore, different approaches must be applied to obtain angle information. The most straightforward technique is using mechanically steered narrow beamwidth elements to perform an angular sweep, like in laser scanners. However, mechanical components are error prone and demand additional control circuitry and thus cost. Therefore they should be avoided in receiver design. Instead, *angle diversity* techniques are utilized in order to exploit the directionality of the receiver. The most common all-electronic techniques are:

- Direction finding using an array of several angularly displaced directional sensors.

[11] provided that the mobile unit is not located on the line joining these two reference units.
[12] also referred to as *direction of arrival* (DOA)

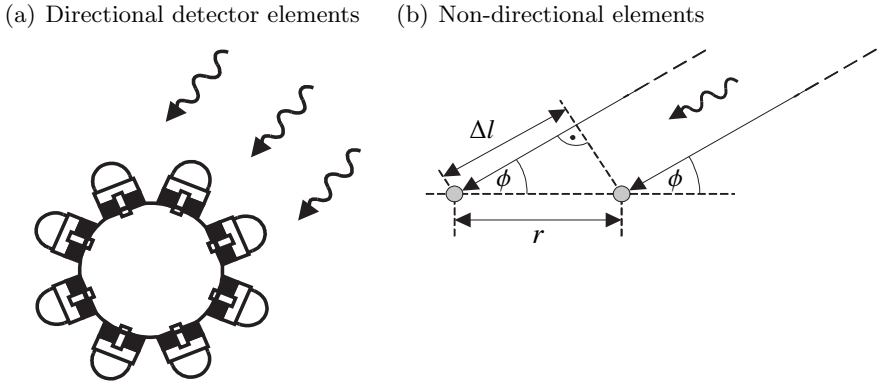


Figure 2.4: Angle of arrival measurement

Fig. 2.4(a) shows such a detector incorporating a number of infrared photodiodes. This design and the process of direction finding is further explained in Ch. 5.

- Direction finding of pulses by TDOA. Fig. 2.4(b) shows an arrangement of two sensors s_1 and s_2 located at a distance r. We assume that both sensors detect a beacon signal originating from the same signal source located at a distance l. For $r \ll l$ we can consider the incoming signals to be parallel. The runtime difference Δt is proportional to the difference of propagation paths, Δl. The angle α can therefore be determined as $\alpha = \arcsin(\Delta l / r)$, where Δl resp. Δt can be determined by a TDOA measurement. Note that in the case of electromagnetic waves, an *extremely* high timing resolution is required for this TOF measurement in order to achieve an acceptable granularity of angles. Therefore this approach is rarely used in practice.

- Direction finding of continuous wave signals by multiple omni-directional detector elements arranged under a fixed displacement (typically, an antenna array). A common direction finding array is the uniform linear array (ULA). In this configuration, the elements are placed on a straight line having equal spacing. Typically, this inter-element spacing is in the order of half the wavelength of the carrier frequency of the received signal. The array usually consists of $4 - 10$ elements. The signal received by either antenna is sampled and passed to a digital signal processor (DSP), where a direction finding algorithm is applied. Often used approaches include beamforming and the MUSIC algorithm [77]. The angular resolution is limited by the number of elements, signal to noise ratio, hardware non-idealities and calibration errors [56].

2.4.8 Comparison of absolute positioning techniques

In the previous sections, trilateration, hyperbolic localization and triangulation have been discussed. Their fundamental difference is that they are based on either TOA, phase shift or

RSS, on TDOA or on AOA measurements. RSS has a very low accuracy and thus strongly depends on additional algorithmic support. TOA is based on the assumption that all units have a synchronized time reference. The same is true for TDOA except for the fact that the time of emission of the signal is irrelevant. Hence, only the reference units, but not the mobile unit, must be synchronized. The advantage of AOA techniques over TOA and TDOA is that no units need be synchronized at all. However, the trade-off is that significantly larger and more complex hardware is required. Furthermore, the location estimate degrades as the mobile target moves farther from the reference units. The discussed techniques are wide-spread among state of the art implementations. In certain systems, hybrid techniques (e.g. TDOA plus AOA) are used. In most systems, redundant reference units are deployed in order to increase system scale and overcome shadowing of single reference units. The resulting scheme is sometimes called *multilateration* to indicate the use of additional units. An overview of existing implementations is given in Sec. 3.

2.5 Stochastic methodologies

An important issue in any sensor system is measurement uncertainty: Measurements are only accurate to a certain degree. As a result, the estimate of the system is erroneous. There are several pieces of information that – if available – can help to delimit this error. Knowledge about the *statistical error distribution* can be exploited to make assumptions of a dynamic system's state. In the case of a localization system, the *state* of the system is defined by 2D or 3D coordinates, but it can as well incorporate orientation and velocity information. Note that the error distributions may be individually shaped for each sensor. It is a challenge to merge the output of different kinds of sensors. This process is commonly referred to as sensor fusion. An assumption on the current location of a mobile target (commonly called *belief*) can further be backed up if information on the *system history* (i.e., the previous state) is available. For instance, it is likely that if a person resides at position (x, y) at one point in time, he/she will still be close to (x, y) one moment later. For this inference, knowledge about the *dynamics of the system* must be available (e.g., how fast can a person move at most).

The described problem has been addressed and mathematically expressed by researchers for many years. Therefore, a large number of approaches are available today. Particularly Bayesian filters have raised much attention, and different variants can be found in uncounted applications. These variants include the Kalman filter and the particle filter, the first of which is known for its computational efficiency using standard matrix computations. A short survey on this topic can be found in the work by Fox et al. [33]. Bar-Shalom et al. provide an in-depth understanding of probabilistic inference in localization applications [5].

2.6 Employed signal technologies

Not only the positioning technique is a characteristic of a localization system, but also the technology it is based on. Absolute localization systems are commonly ultrasonic, infrared, optical, magnetic field or radio based. These technologies differ in many properties, most importantly range, robustness against ambient variations and signal propagation issues. Among the latter are signal velocity and degree of reflection, scattering and attenuation at walls and obstacles. Naturally, the design of the transceivers and the data processing hardware differs according to the signal technology. Therefore, cost, power consumption and the required computing resources can vary significantly.

The primary source of error in indoor localization is multipath propagation caused by signal reflections that overlap with the direct line of sight signal at the receiver. This is true for both sound and electromagnetic waves. The localization techniques described in the previous sections rely on the assumption that the direct line of sight signal can be discriminated from echos resulting from other propagation paths. Therefore, depending on the applied technology, more or less effort must be undertaken in order to detect the signal component that arrives first.

Within a building, the ambient conditions such as temperature and humidity are subject to temporal and spatial variations. Temporal variations take place at different time scales, e.g. the time of day or season of the year. Spatial variations occur due to heaters or air conditioners, or direct sunlight in certain parts of a room. According to the utilized technology, these variations can have a heavy impact on the accuracy of the localization system.

In the following, aspects of the above-mentioned signal technologies are stressed.

- Ultrasound is often used in indoor localization. Due to the low propagation speed highly accurate range measurements can be accomplished without much effort. For the same reason, multipath components can be separated easily. The speed of sound is given approximated by

$$v_{\text{sound, air}} = 331.0\frac{\text{m}}{\text{s}} + 0.6\frac{\text{m}}{\text{s} \cdot {}^{\circ}\text{C}} \cdot T.$$

  The equation shows that the speed of sound is significantly affected by temperature changes. Depending on the desired accuracy, this effect must be combated in the design of a localization system. Contrary to ultrasound, the propagation of electromagnetic waves may be considered robust against ambient changes in air. Another drawback of ultrasound is the fact that the range of these signals is rather limited.

  The most commonly used – and therefore most inexpensive – ultrasound transducers operate at a center frequency of 40 kHz. Contrary to humans who do not hear sound above 20 kHz, some animals do. Some pets are able to hear frequencies up to 50 kHz (dogs) or 60 kHz (cats). Being permanently exposed to a stream of acoustic pulses may be perceived as noise pollution by these animals. Hence, these systems are perhaps not

appropriate for domestic use. Naturally, transducers operating at higher frequencies can be employed. Unfortunately these devices are more expensive. Furthermore, they demand higher transmit power levels because the attenuation of sonic waves in air increases with frequency.

- Infrared components are compact and give rise to very low cost. One of the major challenges is suppressing interfering noise, mainly resulting from ambient light sources. Depending on the wavelength, IR reflects off most surfaces of indoor environments. The typical range is up to 5 m. Usually, due to scattering, only the first reflection must be considered in multipath considerations. Furthermore, equal wavelengths can be used in adjacent rooms without interference because the entire infrared signal is blocked by obstacles or walls. Infrared can roughly be subdivided into active and passive approaches. Active infrared denotes systems where objects or beacons are equipped with IR transmitters such as light emitting diodes. In passive systems, objects are localized as a result of their natural emission of radiation. Of course, radiation is not only emitted by target objects but also by the environment. Therefore, only objects at a temperature other than that of the environment (e.g. humans in a room at 25 °C) can be detected.

  The indoor infrared channel has found good coverage among communication technology researchers in recent years [9, 19]. Although it is not subject of this work, it may be reasonable to add communication capabilities to a localization system for channel arbitration or exchange of collected sensor data.

  IR is well applicable in sensitive areas like hospitals or in environments infested by strong EMI, e.g. industrial production sites. In these environments, radio may be critical since it interferes with vital medical equipment, or operation may be heavily confined due to EMI. Although infrared is immune to such interference, it is of course subject to noise. In most indoor environments there exists significant ambient natural (sunlight) and artificial (incandescent and fluorescent) lighting inducing noise in the infrared receiver [65]. The interference of these noise sources can prevent infrared systems from operating at all or demand excessive transmission power levels. The emitted power is limited by concerns of power consumption and eye safety, especially at wavelengths $< 1400\,\text{nm}$ [3, 47].

- Considerable advances have also been achieved in vision based localization systems. With the increasing sales of digital cameras and pervasive integration of cameras in mobile phones, components like CCD sensors and additional simple optics became inexpensive. However, vision systems still require significant computational and power resources for image processing, feature extraction etc. As a consequence, in spite of relatively inexpensive sensors, cost for vision based localization systems are remarkable, especially when multiple units must be equipped. Furthermore, these systems can be abused for a detailed observation of people[13].

- Using magnetic fields for positioning has the advantage that any nonmagnetic material

[13] In many countries, this fact gives rise to profound doubts on the legality of installing these systems at certain places.

can be penetrated by the field with no loss of position accuracy. These systems utilize sensors placed on the target objects to measure the DC magnetic field generated by a nearby transmitter source. Field propagation is usually limited to about 3 m. Unfortunately this technology is extremely sensitive to interference from a number of sources like CRT monitors. Moreover the field can be affected by nearby objects comprised of magnetic materials. As a result, these localization systems require precise calibration and a static environment. A popular application is motion capturing for computer aided movies and computer games. Due to the apparent drawbacks, this technology is not further considered in this work.

- RF technology can further be split according to the frequency range used, i.e. narrow band and wideband systems. Narrow band systems often used for localization are Wireless LAN (IEEE 802.11b/g), Bluetooth (IEEE 802.15), and RFID (Radio frequency identification). WLAN operates in the 2.4 GHz ISM band at a range of $50 - 100$ m. In recent years, these systems have been installed in thousands of office buildings, homes and public places. Hence it seems worthwhile to employ these infrastructures for localization. Bluetooth has been designed as a cable replacement technology. Like WLAN, it operates in the 2.4 GHz band, but at lower data rate and – depending on the device class – at shorter range. Bluetooth has achieved considerable coverage, particularly in mobile appliances like laptops and PDAs. Neither WLAN nor Bluetooth provide localization capabilities. Commonly, RSS and signal to noise ratio information are exploited to deduce location.

  Also RFID has gained much attention in recent years, particularly in the field of logistics. Within a so-called RFID tag, a small amount of data can be stored. This information can be read and written over an RF link using an RFID reader. RFID tags can be active or passive, that is, they are equipped with a battery or act as a battery-less transponder. Accordingly, their range is limited to a few ten meters for active and $1 - 2$ m for passive tags, respectively. RFID tags are inexpensive and – if passive – need no maintenance like battery charging or replacement. RFID is mostly used to obtain proximity information rather than true localization. That is, we can only tell *that* a tag is within range of an RFID reader, but not *where*.

  Short pulse ultra wideband (UWB) is yet another radio technology suitable for indoor localization. Although being used in military applications for 50 years, UWB has recently gained attention among researchers as well as the industry [69]. It is especially aimed at high speed wireless communication at a short range, e.g. in multimedia applications. Two physical layers for high and moderate data rates (IEEE 802.15.3a/4a) are currently being specified. A wide frequency spectrum can be achieved by classic spread spectrum approaches like DSSS (direct sequence spread spectrum). The alternative idea in UWB is the usage of electromagnetic pulses with an extremely short pulse width in the order of a 1 ns. This approach is particularly advantageous for localization in a cluttered indoor environment, where severe multipath fading occurs. Due to the short pulse length, the direct path can be discriminated from echos more easily. A so called *leading edge detector* that performs this discrimination is presented by Fontana [30].

Besides the standards mentioned here, a large number of proprietary radio technologies are used for localization. RF represents one of the few technologies which does not require direct line of sight condition because radio traverses through many materials. However, the signal is still delayed, scattered, reflected and attenuated on its way, hence affecting the measurement in many ways. As a result, received indoor radio signals are composed of a number of superimposed signal components. If a continuous carrier is considered, these overlays have different phase offsets. As a matter of fact, the signal component to arrive first at the receiver is not necessarily the strongest component. Hence a phase shift measurement can be heavily tampered. It seems therefore reasonable to discriminate the first signal path to enable a time of flight measurement. Unfortunately, discriminating this first path is cumbersome and hence, range measurements have a low accuracy. Likewise, due to multipath fading, the received signal strength (RSS) becomes a bad estimate for the distance between transmitter and receiver.

In some cases, multiple technologies – like radio and ultrasound – are used. This is particularly advantageous for concerns of time measurement and synchronization. In Tab. 2.1, the most important properties of the presented technologies in the context of indoor positioning are summarized. Implementations based on these signal technologies, employing the techniques discussed in the previous sections, are introduced and discussed in Sec. 3.

2.7 Clock synchronization and time measurement in indoor localization

As stated above, for positioning based on TOA measurements, a clock is needed on each of the reference units as well as the mobile target. Furthermore, these clocks must be synchronized. If hyperbolic localization is performed (i.e., TDOAs instead of TOAs are measured), the time of emission at the mobile unit becomes irrelevant. Only the reference units need be synchronized in order to determine the time span between two signal reception times.

In case of radio signals, synchronization of the clocks is critical: even a small synchronization deviation may result in huge measurement errors. A good time measurement accuracy can be obtained using a correlation receiver which searches a maximum between a received and a given waveform. This type of receiver is e.g. used in radar technology to measure roundtrip times. However, in a multipath environment such as the indoor radio channel, this approach suffers from echo signals that interfere with the direct line of sight path. This interference can amplify or attenuate other signal components – an effect called multipath fading. As a result, the received waveform is strongly deformed and the correlation accuracy is degraded.

As an alternative to the correlation receiver, the leading edge of a received waveform can be determined. The resolution of such a TOF measurement system is then determined by the sampling frequency of the detector. Time is actually measured by a digital counter

Signal technology	Assets	Remarks
Ultrasound	Slow signal propagation Facilitates slow clock speed High accuracy	Susceptible to ambient variations Might influence animals
Infrared	Inexpensive components Usually few multipath components	Range: $\leq 5\,\mathrm{m}$ Line of sight required Susceptible to interference from ambient light or thermal radiation sources (depending on the used signal wavelength)
Optical (camera)	No active emitters required	Line of sight required Illuminated environment required Costly signal processing
DC magnetic field	High accuracy	Range: $1 - 3\,\mathrm{m}$ Highly susceptible to environmental interference High deployment cost due to complex calibration procedure
Narrow band radio	Signal traverses through obstacles	Highly susceptible to multipath effects Range of WLAN: $50 - 100\,\mathrm{m}$ Range of Bluetooth: $10 - 100\,\mathrm{m}$ (depending on device class) Range of RFID: $1 - 2\,\mathrm{m}$ (passive tags), a few ten meters (active tags)
Wideband radio	Signal traverses through obstacles High accuracy Fairly robust against multipath effects Low power	Expensive

Table 2.1: Employed signal technologies

clocked at this frequency. E.g., a 100 MHz clock results in $T_s = 10\,\text{ns}$ thresholds. Since the radio signal propagates with the speed of light c_0, an approximate ranging resolution of $c_0 \cdot T_s \approx 3 \cdot 10^8\,\text{m/s} \cdot 10\,\text{ns} = 3\,\text{m}$. This is insufficient for the aimed range of applications. For a resolution of 1 cm, a 3.3 GHz clock is needed. This may require a non-standard technology process and is hence costly. It is however desirable to develop a solution that is applicable to any low cost CMOS process.

A solution to the high frequency problem is the combined employment of different signal technologies. For example, radio signals are used to explicitly synchronize clocks between two units, whereas the actual propagation measurement is based on ultrasound ranging: A (mobile or reference) unit synchronously transmits a radio and an ultrasonic pulse. The receiving unit resets and re-starts its internal clock on reception of the radio pulse and stops it when the ultrasonic pulse arrives. As the time of flight of the radio signal can be neglected compared to the propagation time of the ultrasound wave, the distance can be directly determined from the clock value. Using this method, a good localization precision of a few cm can be achieved [71].

In Chapter 4, a new approach for time measurement and synchronization in hyperbolic localization allowing for the use of RF components is presented. This approach uses relatively low frequency and low cost components but nevertheless enables a high measurement accuracy.

Chapter 3

Related work

Surveys of existing implementations of indoor positioning systems targeting pervasive computing applications have been compiled by Hightower and Boriello [42], Tauber [85] and Muthukrishnan et al. [66]. Some systems for robot localization are summarized by Jiménez [49] and Borenstein, Everett and Feng [13]. The most fundamental implementations and most recent work is presented in this section.

3.1 GNSS based positioning systems

Global Navigation Satellite Systems (GNSS) employ a set of satellites each equipped with synchronized atomic clocks. NAVSTAR, commonly known as the The Global Positioning System (GPS) is the most popular of these systems [45]. Each satellite transmits a radio signal that carries the local clock time and current satellite position, amongst other information. Due to the high precision of the atomic clocks as well as the satellite position, the signal propagation times can be determined very precisely. GPS receivers are equipped with crystal clock oscillators for time measurement. Positioning as well as clock synchronization are accomplished by locking on four code multiplexed satellite signals. With signal processing techniques such as code correlation and integration, signal propagation times with respect to the receiver's clock – so called pseudo ranges – can be determined. When propagation time differences are considered, the local clock bias can be eliminated. A high short term stability for the receiver oscillator is required[1], while the long term stability is of less relevance. GPS cannot be used indoor using standard receivers. The radio signal is attenuated by walls to a degree that it becomes nearly undetectable for these devices. However, two approaches address this problem: The first one, assisted GPS (AGPS), is nowadays largely used in mobile phones. A GPS reference station provides all relevant information (satellite positions, DGPS[2] corrections, Doppler information, etc.) over a mobile phone data link. As a consequence, the AGPS

[1] for the reason of guaranteeing a stable time reference during one measurement

[2] Differential GPS

receiver can more easily lock on the satellite signals. In the second approach, GPS receivers employing large banks of parallel correlators are used to speed up the synchronization with the satellite signals [88]. If the mobile target moves inside a building, positioning is even more complicated. Hence, in practice these approaches are often supplemented with additional sensors such as gyroscopes and accelerometers. Lachapelle carried out several experiments on indoor GPS [58]. The performance of current systems is extremely poor (compared to outdoor GPS): while working fairly well in a wooden building (5 m accurate), tests in a commercial building yielded accuracies from $58 - 75\,\mathrm{m}$. These results clearly show that GPS is currently not qualified for any location aware indoor application.

As an alternative, the GPS concept might be applied for indoor localization, such that beacons act as satellite surrogates – sometimes called pseudolites (see Sec. 3.3) [8].

3.2 Ultrasonic and acoustic positioning systems

Various indoor localization systems have been designed that employ the TOF ranging method using ultrasound and radio discussed in Sec. 2.4.2. An early development is the Bat ultrasonic location system from AT&T which addresses location aware applications in smart buildings [93]. Receivers placed in a square grid under the ceiling, $1.2\,\mathrm{m}$ apart, are interconnected by a serial link. Mobile units subject to localization – termed *Bats* – are small devices that measure $7.5 \times 3.5 \times 1.5\,\mathrm{cm}^3$. They are equipped with an ultrasonic transmitter and a radio transceiver. A Bat that is to be located is addressed via the radio link. It then transmits a pulse of ultrasound at a known time. The times of arrival at three or more beacons is used to compute the Bat's location. The stated system accuracy is $3\,\mathrm{cm}$. However, the fine-grained distribution of beacons makes it difficult to deploy and very costly.

The Cricket location system and its extension, Cricket Compass, developed at the MIT, is a similar implementation [71, 72]. Reported results indicate an average localization accuracy of $6\,\mathrm{cm}$ and an orientation estimate to $3 - 5°$ when beacons are placed every $1 - 1.5\,\mathrm{m}$. However, when the mobile unit is close to a wall, the accuracy decreases significantly due to signal reflections. To keep the system simple, no explicit arbitration among the beacons takes place. Instead, medium access is randomized. A commercial version of the Cricket system is also available at about $300\,\mathrm{US\$}$ per node. One drawback of the system is its dynamic performance. Due to the long update cycle, i.e., the time span until a target has gathered range information towards all nearby reference units, the localization precision decreases with increasing velocity of the target. Probabilistic integration has been added to combat this effect [80]. The relatively large dimensions of the device (approximately the size of a PDA) makes it inconvenient to wear as a badge. Furthermore, the high cost impede a truly pervasive deployment of the system.

The indoor positioning system developed at the University of Bristol can resolve $10 - 20\,\mathrm{cm}$ [73]. The mobile unit is equipped with an RF and ultrasonic receiver. TOF from the reference transmitters installed under the ceiling and the mobile target is measured by an RF

synchronization code in conjunction with an ultrasound sequence. The reference units are deployed in a less fine grid, thus lowering deployment efforts and cost. The system employs trilateration by selecting the three shortest TOF paths.

A general drawback of these approaches is that using standard transducers, ultrasound can usually propagate less than six meters and does not penetrate obstacles. As a consequence a large number of precisely located reference units must be mounted under the ceiling – a scenario which becomes unrealistic in environments such as e.g. large conference rooms. Prohibitively high installation costs can be reduced if deployment strategies are applied often used in distributed sensor networks. Savvides et al. developed the AHLoS architecture (Ad-Hoc Localization System) [75]. This system uses an iterative multilateration strategy to locate a large amount of sensor nodes by using a small number of precisely positioned references. The idea behind iterative multilateration is that a sensor node which can estimate its position using known references becomes a new reference itself. The paper furthermore discusses the ultrasonic/RF ranging approach, but also considers received signal strength for range estimation. While the latter is found to be infeasible for fine-grained resolution, the stated accuracy using ultrasound/RF is 2 cm. However, due to the simple iterative multilateration technique that is applied, the accuracy decreases significantly in large scale networks. A similar system is the DOLPHIN system (Distributed Object Localization system for PHysical-space Inter-Networking) [63]. Depending on the number and deployment of the initial reference nodes, the reported system accuracy is 15 cm – 1 m.

As mentioned in Sec. 2.6, ultrasound is susceptible to ambient variations. This impact can be reduced by a careful selection of reference unit positions and by ambient condition monitoring using temperature sensors. If the temperature variation is only temporal but not spatial then additional reference units can be used to solve for the speed of sound. Another drawback of the ultrasound approach is that the devices must be able to send/receive radio *and* ultrasonic signals which necessitates further hardware (and thus, increases cost). A fact often concealed in the literature is the high directivity of ultrasonic transducers. Hence, the orientation of the devices is of great importance. To reduce this effect, an array of transducers must be used in order to obtain good omni-directional emission. However, this solution increases size and cost of the devices.

The range of ultrasound localization devices is generally low. To decrease the impact of high signal attenuation of ultrasound waves, lower frequencies can be chosen. An acoustic sensing system for robot localization based on the same TOF principle as used in the Bat and Cricket systems is described by Girod and Estrin [37]. The system utilizes standard PC hardware equipped with sound adapters for audio sampling. Synchronization is performed by a wired link using the PC's parallel port. Likewise, the same drawbacks apply as with ultrasonic sensing. Due to the noise generated for ranging, the system is not practicable in an environment where humans work or live.

Bian et al. developed an implicit localization system that estimates sound source locations in a ubiquitous computing environment [12]. The term *implicit* refers to the fact that the system leverages implicit target characteristics (e.g., the human voice), i.e., no active com-

ponents need be worn. In an experimental setup covering three rooms, 16 microphones were deployed, organized in arrays of four microphones. The reported mean accuracy is 13.5 cm with a standard deviation of 27.3 cm, at an update rate of $1-5$ s for continuous human talking.

3.3 Positioning systems based on narrow band RF ranging

Locata is a positioning technology developed at the University of New South Wales, Sydney, Australia [7]. It is meant to extend the capabilities of GPS in urban and indoor environments. For this purpose, ground-based transmitters called *pseudolites*[3] are deployed that emit signals similar to GPS satellite signals. Their signal power is significantly higher compared to the weak GPS signals, hence enabling the indoor use of some standard GPS receiver chip sets[4]. All pseudolites need a precise common time base. The synchronization procedure between two pseudolites *A* and *B* works as follows: *A* generates and transmits a signal which is received by *B* after the physical propagation time. *B* generates its own signal, and adjusts its local oscillator using direct digital synthesis to bring the difference of the two signals to zero. Finally, the oscillator is further tweaked to correct for the geometrical offset between *A* and *B*. This process is performed in a cascade until all remaining pseudolites have been synchronized. Barnes et al. describe an indoor test setup on an industrial site [6]. In a row of experiments, the position of a ceiling mounted crane is observed. The reported mean location accuracy is below 2 cm. This extremely good result could be achieved because depending on the crane position, there existed an undisturbed line of sight between pseudolites and the receiver. In [8], performance of the system is analyzed for a moving target in an office building, with pseudolites being positioned outdoors. In this case, sub-meter precision is reported. Commercialization of the system is currently underway by the Locata Corporation[5]. Although no explicit numbers are presented, it can be assumed that a large scale deployment of Locata, with many mobile targets, results in significant cost: each mobile unit must be equipped with its own GPS-like receiver. Besides being costly, these devices are known for their high power consumption. This fact makes Locata unfavorable in pervasive computing applications, where lightweight tags are worn by people or attached to objects.

PinPoint is a commercial indoor localization system by RF Technologies (formerly PinPoint Corp.[6]) [94]. It is particularly aimed at medical facilities and health care organizations. The infrastructure consists of one or multiple cells, each consisting of a number of antennas and a cell controller. Each cell controller is connected to a host computer. Target objects and patients are equipped with tags. The system uses roundtrip time measurements for ranging, in order to avoid the need for high precision synchronization. Localization is performed by multilateration. At a cell range of 30 m the achieved static accuracy is 1 m. In order to provide a one year lifetime of the tags, their duty cycle is kept to 1/5 s.

[3] from *pseudo-satellites*
[4] with altered firmware, though
[5] www.locatacorp.com
[6] www.pinpointco.com

3.4 Proximity positioning systems

Yet another indoor localization approach is the proximity sensing technique where the position of a target is approximated by selecting the location of the closest reference unit. Accordingly, no location computation is necessary. Alternatively, if multiple reference units are within range, the centroid between these units may yield a better estimate.

The Active Badge indoor localization system pioneered localization in pervasive computing. It is based on proximity information acquired by ceiling-mounted infrared receivers [92]. These devices listen to unique sequences transmitted by badges worn by people at the breast pocket or waist position. Such a sequence is transmitted every 15 s. Unfortunately, the accuracy of this approach is limited to room granularity. Furthermore, the low update rate results in large estimation errors when people move.

An indoor navigation system for augmented reality is presented by Hallaway et al. [39]. High power infrared transmitters are deployed on the walls at regular distances of approximately 2.8 m. These beacons emit unique code sequences in an unsynchronized manner. The receiver is attached on a helmet worn by the user. It consists of eight IrDA dongles that are angularly displaced by 45° each. An elliptical-shaped active area is assigned to each transmitter. Within this area, the minimum signal power required to excite the receiver is exceeded. Therefore, if multiple code sequences are received, the cross section of the respective ellipses yields the location estimate. The positional accuracy of the system is about 1 m for a static user. Unfortunately, the update rate of 1 Hz is rather low, resulting in much higher deviations when the target is moving. A Kalman filter is employed for tracking. Besides movement in the plane, orientation of the user is determined with an accuracy of $5 - 10°$.

A positioning approach using Bluetooth is described by Hallberg et al. [40]. Client location is derived from proximity to one or several Class-3 Bluetooth devices. According to this class, the maximum error is limited to approximately 10 m. If a target moves in the environment, devices get into or out of its range and new links must be established – a procedure which can consume a considerably long time. Accordingly, the update rate is limited by this period. In an experiment, the reported average position retrieval rate was only 19.2 s.

The position of people in a smart room can also be obtained using a sensitive floor. Load cells located in a grid fashion under the floor tiles or carpet yield the varying spatial weight distribution of the floor. An early implementation is the ORL Active Floor by Addlesee et al. [2]. A similar system is presented by Orr and Abowd [70]. In both approaches the sensor grid has a spacing of 50 cm. Positioning accuracy is defined by the dimensions of the floor tiles. Both work concentrate on user identification by analysis of footstep data, primarily based on hidden Markov models. An inexpensive alternative that focuses on location estimation has recently been presented by Kaddoura et al. [51]. The costly load cells used in the older approaches are replaced by flat inexpensive pressure sensors. Furthermore the deployment is slightly altered and the grid spacing enlarged to 60.5 cm, with a dead zone of 9.5 cm between the tiles. A simple model to distinguish between very few types of human walking is presented.

Depending on the granularity of deployed anchor devices, proximity based localization systems yield more or less coarse grained location information. The achievable low resolution limits this method to a very small number of applications, where precise location estimates are unnecessary.

3.5 Positioning systems based on RF RSS ranging

The RADAR system developed by Microsoft Research uses an IEEE 802.11b WLAN infrastructure to perform localization of clients [4]. For this purpose the received signal strength (RSS) as well as the signal to noise ratio (SNR) of received signals are analyzed. Two different approaches were implemented. In the first approach, a map model was set up by RSS and SNR measurements which were performed at various places. The client location was then estimated by searching the nearest neighbor in the signal space. In the second approach, the complex and expensive training phase was substituted by applying a signal propagation model. The reported average localization accuracy was 4.3 m.

Several other research groups have set up similar systems [43, 95, 59] which do not differ significantly from the RADAR approach. A related positioning system is commercially available from the Finnish company Ekahau[7].

A Bluetooth positioning system based on RSS readings of Class-2 devices is described by Feldmann et al. [28]. The reported average accuracy is 2.06 m.

In the LANDMARC architecture, an RFID installation is used to perform localization [67]. However, RFID readers usually do not provide signal strength information. Instead, range is estimated by cycling through eight available power levels. Based on the signal strength received by the reader, it will report or ignore a tag. Power level 1 corresponds to the shortest range and level 8 to the longest range. The maximum range of the readers is 150 m. In return, the tags are equipped with button cell batteries. Besides the target tags which are subject to localization, a number of reference tags must be deployed, separated $1 - 2$ m from one another. From these reference tags, environmental features can be extracted. The assumption is that close-by tags are subject to the same influences, very similar to the idea of differential GPS. In an experimental indoor setup, four readers were placed $4 - 7$ m apart. The readers are interconnected by WLAN links. The reported accuracy is 1 m at a precision of 50%. However, it is stated that the error increases significantly in a dynamic environment, e.g. when a person stands in front of a tag.

[7] www.ekahau.com

3.6 Ultra wideband positioning systems

One of the major issues in indoor position estimation using electromagnetic waves is multipath signal propagation. Fontana reports of an ultra wideband (UWB) based localization system that is capable of detecting the leading edge of a received waveform [30]. Hence, all echo signals can be discriminated from the direct path. The high precision of this detector also enables precise synchronization of the reference units' clocks. Experiments were made in a severe multipath shipboard environment [31]. Results show that UWB is well suitable for ranging even under tough multipath conditions. A commercial variant of this system called PAL650 Precision Asset Location System is available from US based company Multispectral Solutions, Inc.[8] (MSSI) [32]. The stated range is 300 ft (91 m), the accuracy is 1 ft (30 cm). Unfortunately, the update rate is only 1 Hz.

Ubisense, developed at the University of Cambridge, is another commercial UWB based localization system [87][9]. Localization is performed using both TDOA and AOA measurements. To be able to perform the latter, the reference units are equipped with antenna arrays. Currently[10], a demonstration kit including comfortable project planning and CAD software is available at 15.000 €. The guaranteed accuracy is 15 cm. Communication among tags and reference units is accomplished using an additional RF link, not via UWB. Currently, the reference units are synchronized over cable. This means that unnecessarily there exist two additional communication links besides the UWB link. Ubisense currently yields the best accuracy of all RF based indoor localization systems. Contrary to most other systems presented here, it has been proven that it works not only under laboratory conditions, but also in real world installations. However, the high performance is bought for high system cost. Therefore, Ubisense is currently not feasible for pervasive installations.

3.7 Angle of arrival approaches

Krohn et al. describe a small scale system to determine position and orientation of nearby personal appliances (laptops, PDAs, etc.) towards each other [57]. Angles and distances are measured using a circular array of 12 IR transmitters and receivers. No special efforts are undertaken in direction of arrival finding. It is stated that according to the angular characteristics of the IR LED and the photodetectors, the maximum emission and reception angles are 11° and 15°, respectively. The system does not rely on any pre-installed infrastructure. Instead, one device is chosen as the origin of an arbitrary coordinate system. Distances d_{ij} and angles ψ_{ij} between devices i and j are determined. Two strategies are presented to infer locations, a non-linear regression algorithm being the more robust one. Using this algorithm and four devices places in an area of $160 \times 200\,\mathrm{cm}^2$, 95% of the location estimates are accurate to 10 cm in position and 14° in orientation.

[8] www.multispectral.com
[9] www.ubisense.net
[10] 08/2005

3.8 Comparison

In Tab. 3.1, the key parameters of a selection of localization systems are summarized. These implementations have been selected either because they yield high accuracy, were the first to introduce certain new concepts, or are inexpensive. It must be kept in mind that the given figures have been acquired under fundamentally different test conditions. Some figures (e.g. cost) are available for a few systems only. In some cases, cost can be estimated when carefully studying the architecture of a system. For instance, if multiple mobile targets must be equipped with a GPS receiver, parts of the system cost can directly be inferred[11]. However, these assumptions can only be done in rare cases.

[11] Low cost GPS receivers currently are available at around 100€.

System	Approach	Accuracy (and precision)	Scale	Cost	Constraints
Bat	Ultrasound proximity	3 cm	1 cell: 10 m^2	Moderate	Dense sensor grid required
Cricket	Ultrasound/RF TOF	6 cm/3 – 5°	9.14 m w/o any obstacles	300 US$ per node	Dense sensor grid required
Active Badge	IR cellular proximity	Room occupancy	Room scale	Very low cost	Low accuracy, low update rate
ORL Active Floor	Load cell proximity	50 cm	Building scale	High	High installation cost
Locata	RF (TOA)	< 2 cm, < 1 m in cluttered indoor environment	Several 100m	High	Power consuming, high receiver cost
PinPoint	RF (roundtrip time)	1 m	30 m per cell	n/a	Low update rate
RADAR	RSS ranging and SNR analysis	4.3 m (50%)	WLAN range	Low	WLAN adapters needed on tags
MSSI	UWB (TDOA)	30 cm	91 m	n/a	Low update rate
Ubisense	UWB (TDOA and AOA)	15 cm (100%)	1 cell: room size	15.000 € for demo kit	Expensive
GPS	RF TOA	5 m	Global	High	Unavailable indoors
Indoor GPS	RF TOA	5 – 75 m	Global	High	Low accuracy

Table 3.1: Overview of a selection of indoor localization systems

Chapter 4

Implicit clock synchronization and low cost resolution enhancement in hyperbolic localization

In Ch. 2, two key problems of hyperbolic localization – i.e. localization based on time difference of arrival (TDOA) measurements – have been discussed: clock synchronization and time measurement. Solving these problems at sufficient accuracy and low cost yields the major challenge in many indoor applications. In this chapter, low cost solutions for both of these problems are proposed and discussed.

First, a TDOA measurement scheme that implies clock synchronization of the participating units is proposed. Based on the distance differences obtained from this scheme, the location is determined. The approach is mathematically described assuming ideal clocks[1].

The second problem addressed in this chapter is time measurement. This is a particularly critical task when electromagnetic waves are used which propagate at the speed of light. In this case clocks must operate at a very high accuracy and are therefore complex and expensive. A time acquisition approach for indoor hyperbolic localization using a sequence of short pulses is presented. Time measurement is performed by low cost oscillators that operate at a relatively low frequency in the order of 100 MHz which are biased in frequency and subject to jitter. The method is well suited for e.g. ultra wideband electromagnetic waves. The localization approach includes the implicit clock synchronization approach discussed before. I.e., the synchronization which was first shown for ideal clocks is now adapted to unstable clocks that have a low time resolution.

The last aspect of hyperbolic localization covered in this chapter deals with the propagation of the TDOA measurement errors during the location computation. It is shown that the

[1]Ideal in this sense means that clocks run exactly at their nominal frequency and do not drift. Furthermore, these clocks have an infinitely high time resolution.

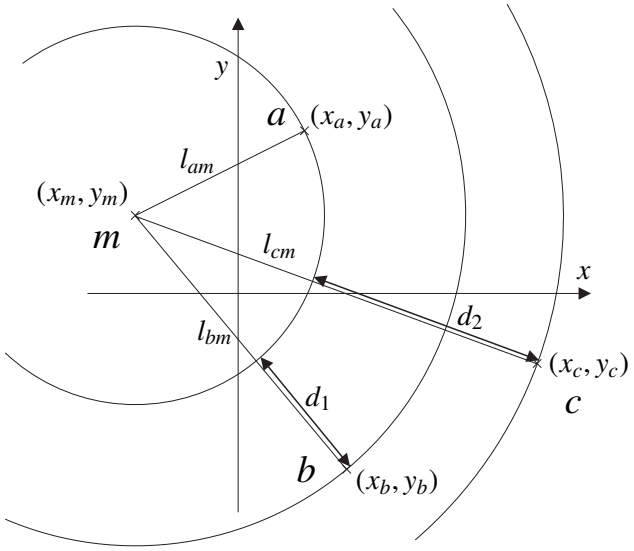


Figure 4.1: Hyperbolic localization setup

accuracy of hyperbolic localization depends on the constellation of reference units and the location of the mobile unit that is to be located. A slim closed-form estimation of the localization accuracy is presented. This helps to configure the physical position of the reference units in order to obtain a desired accuracy level.

4.1 Localization system setup

Hyperbolic localization with implicit clock synchronization works as follows: For localization in the plane, at least three reference units a, b and c are deployed in the environment. The positions of these units must be known and are denoted (x_a, y_a), (x_b, y_b) and (x_c, y_c) in the following. Each reference unit must be equipped with a transmitter and a receiver module that are used for localization and preferably also communication. The mobile unit is only equipped with a transmitter. Furthermore, every reference unit has a local clock. For the moment, we assume that all local clocks are running at the same constant speed and are not synchronized with each other. Furthermore we assume that the clocks have an infinite temporal resolution. At a later point, we will introduce different clock speeds, low temporal resolution and clock jitter. The signal propagation speed v is known while the position (x_m, y_m) of the mobile unit m is to be determined. This setup is illustrated in Fig. 4.1.

Prior to localization, two reference measurements are required in order to determine the temporal relation between the reference units. These measurements will be explained in detail in Sec. 4.2. In order to determine the position of the mobile unit m, only a single beacon signal transmitted from m is required. As soon as a reference unit receives this signal, it takes a snapshot of its current clock value. From these values determined by units a, b and

c the time differences of arrival can be determined. Based on these values, the position can geometrically be computed. This position computation for the mobile unit is described in detail in the literature (e.g. [62]) and has been discussed in Sec. 2.4.5.

This work restricts to a simple setting of one mobile unit and three reference units. Due to the small number of reference units, single measurement errors or signal loss may have a heavy impact on the localization precision. To overcome this problem, an over-constrained set of equations and stochastic methodologies are commonly used [33]. Fortunately, the approach presented here can be extended to a larger amount of reference units to perform maximum likelihood estimation [21]. Besides, this modification can also extend the system to a larger scale. However, in many applications, system cost is a major concern. Hence it is reasonable to analyze the system at minimum requirements.

4.2 Implicit clock synchronization

TDOA measurements lead to distance differences between the mobile unit and the reference units. The distance difference d_1 is defined as the difference of the direct signal propagation paths from the mobile unit m to reference unit a and b, respectively. Accordingly, d_2 is the difference of the direct signal paths from m to a and m to c:

$$d_1 = l_{am} - l_{bm} \tag{4.1}$$

$$d_2 = l_{am} - l_{cm} \tag{4.2}$$

To explain how the distance differences d_1 and d_2 are obtained from the arrangement, additional parameters ($p, q \in \{a, b, c, m\}$) are introduced:

- l_{pq}: Physical distance between units p and q $\quad (l_{pq} \equiv l_{qp})$
- t_{pq}: Signal propagation time between units p and q $\quad (t_{pq} \equiv t_{qp})$
- t_p: Local time of unit p
- Δt_p: Clock offset between global and local time of unit p
- $t_{s,p}$: Local signal transmission time at unit p
- $t_{r,pq}$: Local signal reception time of a signal transmitted from unit p to unit q

The local times t_a, t_b, t_c and t_m are interrelated to each other with regard to a *virtual global time* t_g and their respective clock offsets[2]:

$$t_g = t_a + \Delta t_a = t_b + \Delta t_b = t_c + \Delta t_c = t_m + \Delta t_m$$

The signal propagation times between m and the reference units can be determined using the following dependencies:

$$t_{pq} = \frac{l_{pq}}{v}, \quad p, q \in \{a, b, c, m\} \tag{4.3}$$

[2]Note that for the moment, we assume Δt_p to be constant but unknown. Real worlds issues concerning the constance of Δt_p will be discussed later in this section.

v is the signal propagation speed.

In the following, three measurements are discussed that are required for localization: The actual localization measurement, as well as the two already mentioned reference measurements.

1. Localization measurement
   The mobile unit m emits the signal at its local time $t_{s,m}$. It is received by the reference units a, b and c at their local times $t_{r,ma}$, $t_{r,mb}$ resp. $t_{r,mc}$. With consideration of the units' clock offsets Δt_a, Δt_b, Δt_c and Δt_m, three equations can be set up[3]:

$$t_{am} + t_{s,m} + \Delta t_m = t_{r,ma} + \Delta t_a \tag{4.4}$$
$$t_{bm} + t_{s,m} + \Delta t_m = t_{r,mb} + \Delta t_b \tag{4.5}$$
$$t_{cm} + t_{s,m} + \Delta t_m = t_{r,mc} + \Delta t_c \tag{4.6}$$

   When we subtract Eq. 4.4 and 4.5, as well as Eq. 4.4 and 4.6, and apply Eq. 4.3, the distance differences d_1 and d_2 as defined in Eq. 4.1 and 4.2 can be computed as follows:

$$d_1 = (t_{r,ma} + \Delta t_a - t_{r,mb} - \Delta t_b) \cdot v \tag{4.7}$$
$$d_2 = (t_{r,ma} + \Delta t_a - t_{r,mc} - \Delta t_c) \cdot v \tag{4.8}$$

2. Reference measurements
   In order to eliminate the dependencies of the global time, Δt_a, Δt_b and Δt_c, two reference measurements are necessary. For this purpose the reference units b resp. c act as a sender and the remaining two reference units as receivers. These measurements can be regarded as an implicit synchronization of the reference units using the geometric relationship.

   (a) Measurement b (b is sender, a and c are receivers):

$$\begin{aligned} t_{ab} + t_{s,b} + \Delta t_b &= t_{r,ba} + \Delta t_a \\ t_{bc} + t_{s,b} + \Delta t_b &= t_{r,bc} + \Delta t_c \end{aligned}$$
$$\Downarrow$$
$$\Delta t_a - \Delta t_c = t_{ab} - t_{bc} - t_{r,ba} + t_{r,bc} \tag{4.9}$$

   (b) Measurement c (c is sender, a and b are receivers):

$$\begin{aligned} t_{ac} + t_{s,c} + \Delta t_c &= t_{r,ca} + \Delta t_a \\ t_{bc} + t_{s,c} + \Delta t_c &= t_{r,cb} + \Delta t_b \end{aligned}$$
$$\Downarrow$$
$$\Delta t_a - \Delta t_b = t_{ac} - t_{bc} - t_{r,ca} + t_{r,cb} \tag{4.10}$$

[3]Latencies of the transmitter and receiver units are considered constant here and are implicitly included in the clock offsets.

Inserting Eq. 4.9 and 4.10 into Eq. 4.8 and 4.7, resp., yields:

$$d_1 = (t_{r,ma} - t_{r,mb} + t_{ac} - t_{bc} - t_{r,ca} + t_{r,cb}) \cdot v \quad (4.11)$$

$$d_2 = (t_{r,ma} - t_{r,mc} + t_{ab} - t_{bc} - t_{r,ba} + t_{r,bc}) \cdot v \quad (4.12)$$

The signal propagation times t_{ab}, t_{ac} and t_{bc} can be computed out of the geometrical arrangement of the reference units. The clock values $t_{r,ma}$, $t_{r,mb}$ and $t_{r,mc}$ are acquired in the localization measurement. $t_{r,ba}$, $t_{r,bc}$, $t_{r,ca}$ and $t_{r,cb}$ are obtained from the reference measurements and – in a theoretical setting – need be determined only once.

As noted before it was assumed that all clock offsets Δt_i are constant and do not shift in time. This is the ideal case but does not apply for the real world. Time measurements are typically implemented by a counter clocked by a temperature compensated quartz crystal oscillator. However, these devices have limited accuracy due to various noise effects [41]. As a result, the reference units' clocks will drift apart. Depending on the desired measurement precision and the accuracy of the used oscillator, it is necessary to repeat the reference measurements at certain intervals in order to recalibrate the system. Contrary to the reference units, the mobile unit need only be equipped with a relatively inaccurate oscillator since the exact time of signal emission is irrelevant for TDOA. Hence this oscillator can be very low cost.

With Eq. 4.11 and 4.12 d_1 and d_2 can be determined without any explicit synchronization of the units' local clocks. Any initial local clock offset will inherently be compensated. Following the direct calculation method depicted in App. B, the solution to the position computation from d_1 and d_2 is given by a quadratic equation. When evaluating the equation, two situations may occur: The position can be unambiguously or ambiguously determined. I.e., due to the quadratic relationship, two potential positions (x_{m1}, y_{m1}) and (x_{m2}, y_{m2}) are obtained from a pair of d_1 and d_2. Fig. 4.2 shows a plot visualizing the unambiguously and ambiguously computable areas for a room of $6.5 \times 8\,\mathrm{m}^2$. In certain cases the true position can be unambiguously found, i.e. there exists a unique, real solution to the quadratic equation. However, for other values of d_1 and d_2, the equation can yield *two* real solutions. In this case the true position of the mobile unit cannot distinctly be determined. The common idea to solve this problem is to deploy an additional reference unit such that any ambiguities are eliminated. However, the correct position may also be found by plausibility analysis, i.e. by examining whether one of the computed points is located within the boundaries of the room where localization takes place. This of course implies that the outline of the room is known.

Besides the two cases illustrated in Fig. 4.2, another situation may occur. If measurements d_1 and d_2 are erroneous, only complex solutions may exist to the quadratic equation. I.e., no real position (x_m, y_m) can be found. In this case, the direct computation method fails. It is reasonable to discard and repeat the measurement. Alternatively, to make the system more robust against these situations, additional reference units can be deployed [21].

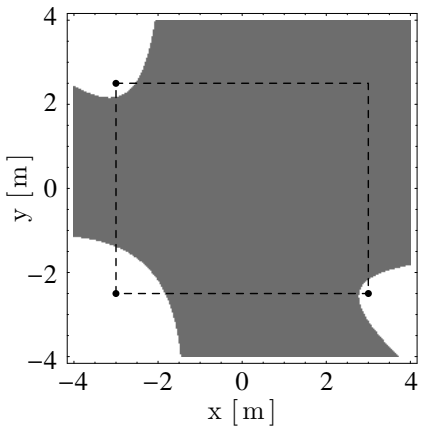


Figure 4.2: Unambiguously (dark) vs. ambiguously (white) computable area. The three dots indicate the reference units, the dashed rectangle can be considered as the outline of a room.

In this section, the physical arrangement of TDOA measurements was depicted. An approach for implicit clock synchronization that yields the distance differences d_1 and d_2 was described. It was also pointed out that for a measured pair of distance differences, the position of the mobile unit can unambiguously or ambiguously be determined. Furthermore, in the case of measurement errors, no solution may exist.

The advantages of the implicit synchronization approach are:

- The local clocks running at each of the reference units do not have to be externally synchronized. This is especially advantageous if radio signals are used to measure propagation delay. Instead of external synchronization, two reference measurements among the reference units are performed. This will implicitly synchronize the reference units.
- The mobile unit must only be able to emit a beacon signal. I.e., it need not be equipped with a receiver. Therefore, the method is well suited for applications where the mobile unit must be low cost and thus requires a low complexity.
- The method is applicable to radio, IR and ultrasound based systems.
- In case of RF and IR, the transmitter/receiver equipment may be used for localization as well as for data transfer. Hence, no special transmitter/receiver modules dedicated only to localization are required.

4.3 Resolution enhancement by time averaging

In the previous section, it was assumed that the local clocks residing in the reference units are ideal. Unfortunately, in the real world this is not the case. In order to achieve the

desired accuracy, the clocks must be able to resolve differential times of arrival to a sufficient level. The achievable resolution of an TDOA measurement strongly depends on the signal propagation speed. As discussed in Sec. 2.7, this is particularly critical when electromagnetic waves such as radio are used. Unfortunately, high speed clock oscillators in the GHz range are expensive. Here, a solution is presented that allows for inexpensive COTS oscillators in the order of 100 MHz. A high resolution is achieved by repeated measurements and time averaging, following a certain scheme.

In order to enhance the resolution of the clocks embedded in the reference units, the TDOA measuring procedure described above is modified as follows.

- As stated above, instead of a single pulse, a transmitter unit p now emits a *contiguous sequence of N_p pulses* with constant frequency $1/T_p$ (see Fig. 4.3). The transmitter is reset at $t_p = 0$. After a delay of $t_{s0,p}$, the emission of pulses is initiated. The resulting signal is denoted s_p.
- The reference units sample the received signal s_{pq} using a clock running at $1/T_q$ (with T_q being the clock period at reference unit q). Furthermore, with each clock tick of T_q a counter $C_{q,p}$ is incremented by one. The indices in $C_{q,p}$ indicate that pulses transmitted by unit p are measured by unit q.

  If a newly sampled value is equal to 1 while the preceding sampled value was 0, the current counter value $C_{q,pi}$ is added to a sum S'_{pq}, with $i \in \{0, 1 \dots N_p - 1\}$. E.g., in the example shown in Fig. 4.3 the counter values $C_{q,p0} = 5$ and $C_{q,p1} = 8$ are added. The sum is initialized to 0 before the measurement begins[4]. The process continues until N_p counter values have been added to S'_{pq}.

Note that the initial value of the counter is non-relevant as long as the counter does not overrun during operation. This can be ensured by resetting the counter after the N_p-th sample. Furthermore, in Fig. 4.3 the transmitter p has a significantly lower frequency than the receiver q. This is not a requirement of the measurement approach. Here, the low frequency was only chosen for an easier understanding. If the transmitter has a higher frequency than shown in the figure, it is reasonable to emit pulses every n_p clock periods T_p, where $n_p \in \mathbb{N}$.

The cumulative time value S_{pq} corresponding to the sum S'_{pq} can be obtained as follows:

$$S_{pq} = T_q \cdot S'_{pq} = T_q \cdot \sum_{i=0}^{N_p-1} C_{q,pi}, \tag{4.13}$$

where $C_{q,pi}$ represents the counter value after the $(i+1)$-th detection of a rising edge of s_{pq}. $C_{q,pi}$ can be replaced such that

$$S_{pq} = \sum_{i=0}^{N_p-1} T_q \left\lceil \frac{i \cdot T_p + t_{pq} + t_{s0,p} + \Delta t_p - \Delta t_q}{T_q} \right\rceil, \tag{4.14}$$

[4]Note the difference between *sum* and *counter*.

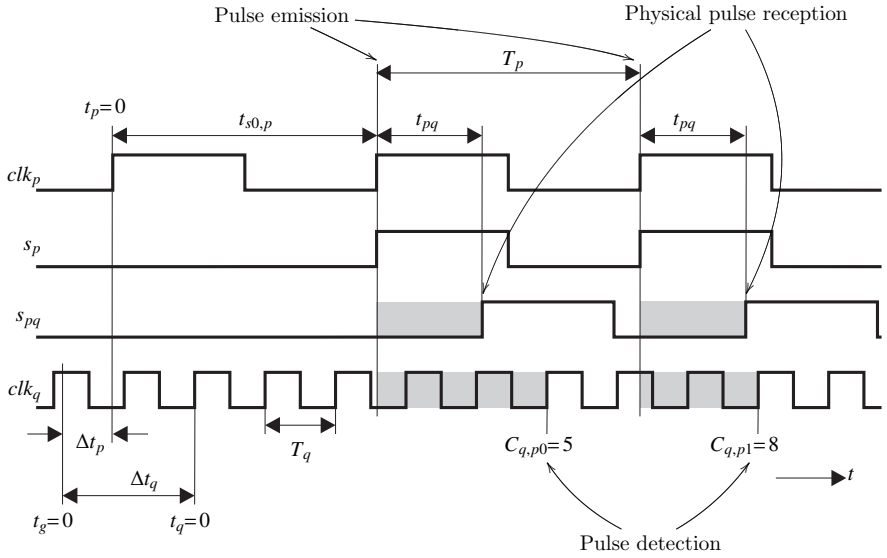


Figure 4.3: Pulse transmission clock scheme. The received pulse is only detected at the next rising edge of the receiver's local clock. For simplicity, in this example, the counter is reset at $t_q = 0$.

where t_{pq} is the time of flight of the signal from unit p to unit q. Δt_p and Δt_q are the clock offsets of units p and q to the *virtual* global clock t_g.

Unfortunately, t_{pq} is argument to a ceiling function in Eq. 4.14 and hence cannot be directly extracted from the equation. However, S_{pq} can be reformulated if the transmitter and receiver clock frequencies are sufficiently "*different*". That is, the phase configuration of the clocks may recur only after a relatively long time. This requirement can mathematically be expressed by

$$k_p T_p \neq k_q T_q, \quad k_p, k_q \in \mathbb{N},$$

where k_p and k_q are no integer multiples of each other and are large (in the order of 100)[5]. Under this condition, S_{pq} becomes

$$S_{pq} = \sum_{i=0}^{N_p-1} \left(i \cdot T_p + t_{pq} + t_{s0,p} + \Delta t_p - \Delta t_q + \Phi \right), \tag{4.15}$$

where Φ is a random variable representing a random process that generates equally distributed real numbers from the interval $[0, T_q)$, yielding a mean of $T_q/2$ and a variance of ${T_q}^2/12$.

[5]In the real world, clocks are always subject to jitter. This fact supports the requirement for "different" clocks.

Averaging and TDOA extraction

In the next step, the average measured time $\hat{S}_{pq}$ is determined by dividing by N_p. For large N_p, the result is described by the central limit theorem [16]: $\hat{S}_{pq}$ in turn is a random variable that is normally distributed (indicated by $\mathcal{N}$) with mean value μ and variance σ^2.

$$\hat{S}_{pq} \sim \mathcal{N}\left(\mu = t_{pq} + t_{s0,p} + \Delta t_p - \Delta t_q + \frac{(N_p-1)T_p + T_q}{2}, \quad \sigma^2 = \frac{{T_q}^2}{12N_p}\right) \tag{4.16}$$

As a result, in comparison to a single measurement, the standard deviation σ of the random process $\hat{S}_{pq}$ decreases by $\sqrt{N_p}$. I.e., the measurement accuracy can be improved by increasing N_p.

Actually, we are not interested in the mean value of $\hat{S}_{pq}$. Instead, we determine the difference of $\hat{S}$ values of two receivers q and r, since they yield the time differences of arrival. Calculating $\hat{S}_{pq} - \hat{S}_{pr}$ yields

$$\hat{S}_{pq} - \hat{S}_{pr} \sim \mathcal{N}\left(\mu = t_{pq} - t_{pr} + \Delta t_r - \Delta t_q + \frac{T_q - T_r}{2}, \quad \sigma^2 = \frac{T_q^2 + T_r^2}{12N_p}\right), \tag{4.17}$$

Note that $\Delta t_r - \Delta t_q$ can be eliminated by the reference measurements described in 4.2. Hence, the TDOA $t_{pq} - t_{pr}$, a term proportional to one of the distance differences shown in Fig. 4.1, can be obtained. As an example, this procedure is shown for $t_{ma} - t_{mb}$:

From Eq. 4.17 and with $E\left(\hat{S}_p\right)$ representing the expected value (i.e., the mean μ) of the random variable $\hat{S}_p$, we get:

$$\begin{aligned} E\left(\hat{S}_{ma}\right) - E\left(\hat{S}_{mb}\right) &= t_{ma} - t_{mb} + \Delta t_b - \Delta t_a + \frac{T_a - T_b}{2}, \quad \text{or} \\ t_{ma} - t_{mb} &= E\left(\hat{S}_{ma}\right) - E\left(\hat{S}_{mb}\right) + \Delta t_a - \Delta t_b + \frac{T_b - T_a}{2}. \end{aligned} \tag{4.18}$$

As described in Sec. 4.2, we can eliminate the clock offsets Δt_b and Δt_a using reference measurement c. For this measurement, Eq. 4.17 yields

$$\begin{aligned} E\left(\hat{S}_{ca}\right) - E\left(\hat{S}_{cb}\right) &= t_{ca} - t_{cb} + \Delta t_b - \Delta t_a + \frac{T_a - T_b}{2}, \quad \text{or} \\ \Delta t_a - \Delta t_b &= E\left(\hat{S}_{cb}\right) - E\left(\hat{S}_{ca}\right) + t_{ca} - t_{cb} + \frac{T_a - T_b}{2}. \end{aligned} \tag{4.19}$$

Inserting Eq. 4.19 into 4.18 finally puts up

$$t_{ma} - t_{mb} = E\left(\hat{S}_{ma}\right) - E\left(\hat{S}_{mb}\right) - E\left(\hat{S}_{ca}\right) + E\left(\hat{S}_{cb}\right) + t_{ca} - t_{cb}. \tag{4.20}$$

t_{ca} and t_{cb} are again determined from the geometric arrangement of the reference units. An analog equation can be found for the TDOA $t_{ma} - t_{mc}$. Note that the result depicted in Eq. 4.20 represents the *expected value*.

However, when the measurement procedure is performed in an experimental setting, the *sample means* $\langle x_{ma}\rangle$, $\langle x_{mb}\rangle$, $\langle x_{mc}\rangle$, $\langle x_{ba}\rangle$, $\langle x_{bc}\rangle$, $\langle x_{ca}\rangle$ and $\langle x_{cb}\rangle$ are obtained[6]. Depending on a number of factors (clock frequencies, number of measurements, etc.), these values will differ from the expected values. In order to account for this fact, we introduce the *measured* distance differences D_1 and D_2:

$$\begin{aligned} D_1 &= v \cdot (\langle x_{ma}\rangle - \langle x_{mb}\rangle - \langle x_{ca}\rangle + \langle x_{cb}\rangle + t_{ca} - t_{cb}) \\ D_2 &= v \cdot (\langle x_{ma}\rangle - \langle x_{mc}\rangle - \langle x_{ba}\rangle + \langle x_{bc}\rangle + t_{ba} - t_{bc}) \end{aligned}$$

For a large number of measurements, N_p, the sample means can be expressed by the series of counter snapshots (see Eq. 4.13):

$$\langle x_{pq}\rangle = \frac{1}{N_p}\sum_{i=0}^{N_p-1} C_{q,pi}, \quad p \in \{b,c,m\},\ q \in \{a,b,c\}$$

In order to judge the outcome of a measurement using the averaging approach, we introduce the measurement errors Δd_1 and Δd_2:

$$\Delta d_1 = |D_1 - d_1|, \quad \Delta d_2 = |D_2 - d_2|$$

Oscillator considerations

In the ideal case, the clocks T_p and T_q are running at a constant fixed frequency. The aim of this work is to use inexpensive COTS (commercial off-the-shelf) quartz clock oscillators for time measurements and general operation of the participating units. In our application, we must consider frequency and phase stability issues. Typically, these effects can be divided in long and short term influences [41]:

- Long term frequency stability is affected by device aging due to internal changes in the oscillator, e.g. the relaxation of stresses. As a result, its frequency will slowly drift from the nominal value. Typical values for aging are 3 ppm (parts per million) in the first year and 1 ppm/ yr beyond. Besides aging, temperature changes can affect the frequency of clock oscillators. In Sec. 4.5 it is discussed how long term bias can be compensated.

- Short term stability is affected by operating temperature, variation of supply voltage, load condition, shock and vibration. Typical values for commercial devices are 10 ... 100 ppm. Phase and amplitude noise as a result of these effects lead to different sorts of jitter. A common definition of jitter can be found in [55]: "*The deviation from the ideal timing of an event.* [...] *Jitter is composed of both deterministic and Gaussian (random) content.*"

[6] instead of the expected values $E(\hat{S}_{ma})$, $E(\hat{S}_{mb})$, $E(\hat{S}_{mc})$, $E(\hat{S}_{ba})$, $E(\hat{S}_{bc})$, $E(\hat{S}_{ca})$ and $E(\hat{S}_{cb})$

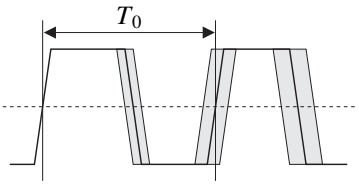


Figure 4.4: Period jitter

Clock jitter

Deterministic jitter of clock signals has a non-Gaussian probability density function and is always bounded in amplitude. It is caused by crosstalk effects from EMI (electromagnetic interference) sources such as adjacent lines and power supply noise.

Random jitter is characterized by a Gaussian distribution. It is caused by a number of stochastic processes including thermal vibrations. The most common jitter measure is period jitter (see Fig. 4.4). It is the time difference between a measured cycle period and the ideal cycle period T_0 and is given as either peak-to-peak or root mean square (RMS). For typical crystal clocks, the period jitter may range from 5 to 50 ps RMS. As we are actually interested in achieving a precision of the same order as these values (a propagation time of 50 ps corresponds to a 15 mm distance in air), clock jitter must be taken into account. Accumulated random jitter is unbounded and can therefore severely affect the measurement precision of the system.

The error sources of deterministic jitter are implementation dependant. Furthermore, this type of jitter is bounded and has hence much less impact than random jitter. Therefore it is neglected in the following. To account for the effect of clock jitter z on $\hat{S}_q$, we model the clock jitter as white Gaussian noise

$$z \sim \mathcal{N}\left(\mu_z = 0, \quad {\sigma_z}^2 = {\sigma_q}^2\right),$$

where σ_q is the standard deviation of the receiver clock of unit q. Of course, the transmitter clock is subject to jitter, too. However, since time *difference* measurements of two receivers are carried out, the impact of the transmitter clock deviations are inherently eliminated.

Finally, under consideration of jitter, the difference $\hat{S}_q - \hat{S}_r$ becomes

$$\hat{S}_q - \hat{S}_r \sim \mathcal{N}\left(\mu = t_{pq} - t_{pr} + \Delta t_r - \Delta t_q + \frac{T_q - T_r}{2}, \quad \sigma^2 = \frac{T_q^2 + T_r^2}{12 N_p} + \frac{2\sigma_z^2}{N_p}\right). \tag{4.21}$$

However, if the period jitter, σ_z, and number of measurements, N_p, are large, then this model is invalid. *Large* in this sense means that the clock jitter may accumulate rapidly, as illustrated in Fig. 4.5: Two clocks are shown in this figure, a nominal receiver clock (top) and a deviated receiver clock which is subject to jitter (bottom). Furthermore, the respective values of the time measurement counters driven by these clocks are shown. The clock bias is

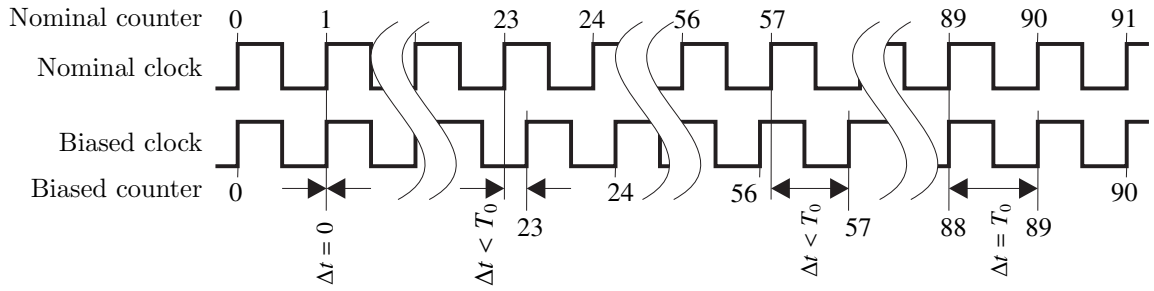


Figure 4.5: Accumulation of clock jitter

illustrated for four points in time. At first, the clocks are in phase and the counters show equal values. The phase offset is denoted Δt, which is zero in this case. In this example, we assume that most of the time, the jitter shifts the clock phase into the same direction. As a result, the phase of the lower clock slowly shifts with time. In the second and third point in time that we consider, the phase offset has grown ($\Delta t < T_0$, where T_0 is the nominal clock period). Finally, in the fourth point in time, the lower clock has shifted by a full clock period ($\Delta t = T_0$). The counter value is deviated exactly by 1. Once the jitter has accumulated into one direction or the other, not a *single* but *many* counter snapshots are biased. In other words, if the jitter has affected the i-th measurement, it is highly likely that it will also affect the $(i+1)$-th measurement. This case is not accounted for in Eq. 4.21. An analytical model that copes with cases like the one illustrated above will become rather complex. An unboundedly growing counter bias might be prevented by resetting the counter after every single measurement[7]. However, alternatively, the proposed measuring method can also be analyzed by simulation. This approach has the significant advantage over the analytical model that the measurement procedure can more easily be modified. As will be shown, a modification of this procedure is necessary to delimit the impact of jitter.

4.4 Simulation results

A simulation model of the localization system based on the time averaging approach described above has been designed. The model was implemented using VHDL. In the first set of simulations, it was evaluated how fast the measurement errors Δd_1 and Δd_2 decline, depending on various configurations of clock frequencies and jitter. Data acquisition was done following a certain pattern, shown in Fig. 4.6. The gray boxes represent time slots in which the actual measurements are performed, as depicted in Fig. 4.3.

One such slot is referred to as transmission mode b, c or m. The measurement procedure is therefore:

- b is sender, a and c are receivers. N_b measurements are taken and averaged.

[7]In this case, the formulas given above must slightly be modified.

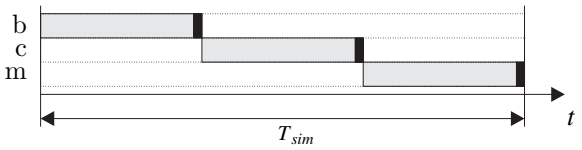


Figure 4.6: Transmission pattern for one-time synchronization

- c is sender, a and b are receivers. N_c measurements are taken and averaged.
- m is sender, a, b and c are receivers. N_m measurements are taken and averaged.
- d_1 and d_2 are computed according to Sec. 4.2 and 4.3.

Note that the time slot reserved for each transmission mode is fix. Hence, if all clock frequencies are different, $N_b \neq N_c \neq N_m$ holds. That is, the number of measurements that can be performed in the fix time slot depends on the frequency of the current transmitter. Another aspect that has not been mentioned so far is the following: In practice, mode switches do not occur instantly, but take a certain time[8]. This time span is referred to as *switch delay* t_{SD} and is indicated in Fig. 4.6 by black boxes. During this time, no measurements are performed. The switch delay can generally be considered as a guard interval that prevents one transmitter's pulses to be mistaken as another transmitter's pulses.

Fig. 4.7 shows measurement errors Δd_1 and Δd_2 against measurement time for two different device configurations. In the first case, the clocks of the devices a, b, c and m run at very similar speed. In the second case, the clock periods vary significantly. The clock jitter of all clocks is 50 ps RMS in all cases.

The first observation to be made is that in fact, the measurement error decreases with rising simulation time. Also keep in mind that without the approach introduced here, a 100 MHz clock can resolve only 3 m. Simulation results (a) and (b) show that stronger differing clock frequencies yield a faster decline of the error. It can also be seen that although declining rapidly in the beginning (until 15 µs), the error does not converge to zero in the long term (Fig. 4.7(b)). The reason for this result is the accumulated jitter. In the experiments to follow, clock configurations are chosen as in Fig. 4.7(b).

Periodic resynchronization

As indicated above, accumulated jitter can corrupt the entire measurement: Once the receiver clock phase (as a result of accumulated jitter) has shifted by a certain amount, not only the *current* but also *all subsequent* measurements will be affected. As a consequence, the overall result is heavily biased. This implication can be limited by periodic resynchronization: Instead of one long burst of measurements, many smaller bursts are performed. Every time all modes

[8]For instance, a certain time might be needed until the power amplifier of the transmitter is operational, when switching a unit from receive to transmit mode.

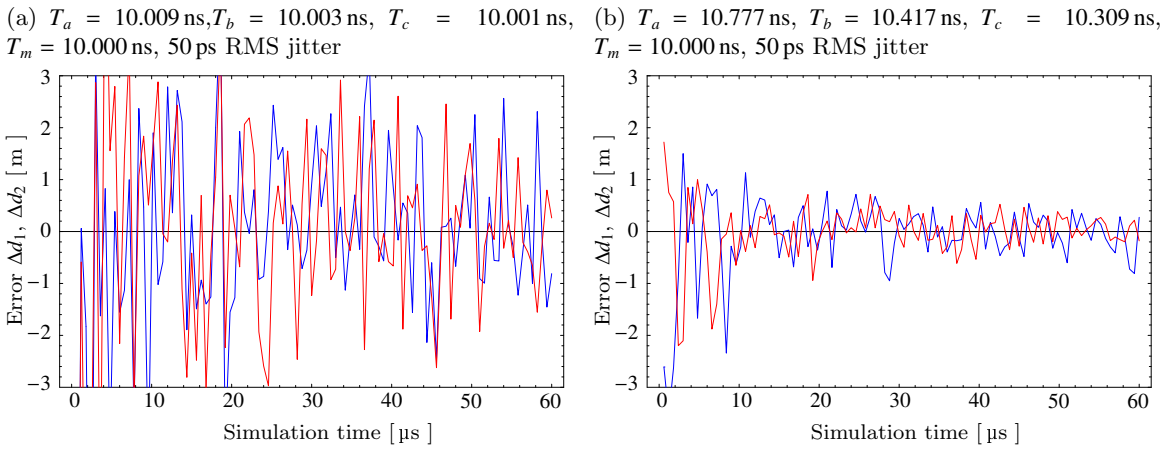


Figure 4.7: Measurement errors Δd_1 (blue) and Δd_2 (red) vs. measurement time, without periodic clock synchronization. Stronger differing clock frequencies yield a faster decline of the measurement error.

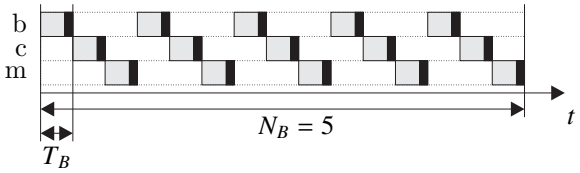


Figure 4.8: Transmission pattern for periodic synchronization

m, b and c have been passed through, the system has been implicitly resynchronized. Hence, jitter cannot grow unboundedly. A modified transmission pattern following this idea is shown for $N_B = 5$ burst cycles in Fig. 4.8. A burst cycle is a set of three bursts, each of which represents one of the three transmission modes. The time span of one burst is denoted T_B. Instead *one* data set of size N_b, N_c and N_m, respectively, *multiple* data sets of size $N_{b,j}$, $N_{c,j}$, and $N_{m,j}$ are now obtained, with $j \in \{1, \ldots, N_B\}$. To retrieve the overall number of measurements in each mode, these must simply be summed up:

$$N_b = \sum_{j=1}^{N_B} N_{b,j}, \quad N_c = \sum_{j=1}^{N_B} N_{c,j}, \quad N_m = \sum_{j=1}^{N_B} N_{m,j}.$$

To obtain the time differences of arrival, D_1 and D_2, only a few changes have to be considered. The pulse transmission scheme for the modified, resynchronizing measurement procedure is shown in Fig. 4.9 (compare Fig. 4.3). The figure depicts an excerpt of the transmission pattern (Fig. 4.8), where at first mode m is entered, followed by modes b and c. Then again, mode m is entered and the sequence is repeated. According to these modes, units m, b and c, respectively, act as transmitters. In the figure, the time measurement is shown at unit a. The pulses emitted by m, b and c are received at unit a after the respective physical propagation time. The received signals are denoted s_{ma}, s_{ba} and s_{ca}, respectively. The

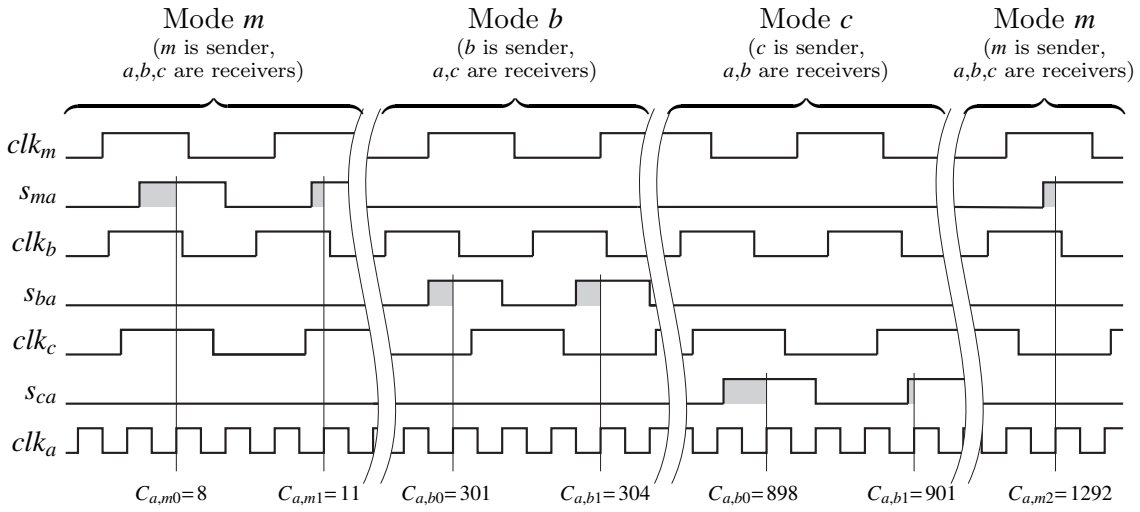


Figure 4.9: Pulse transmission clock scheme with resynchronization

pulses are *detected* at the rising edge of the receiver clock, clk_a. When a pulse is detected, a counter snapshot is taken according to the current transmission mode: In mode m the counter snapshots $C_{a,mi}$ are taken, in mode b the snapshots $C_{a,bi}$, and in mode c the snapshots $C_{a,ci}$, where i represents the number of the snapshot taken in the respective mode. Note that the counter is never reset[9]. As a result, very large *gaps* in the counter snapshots can be observed, e.g. between the values $C_{a,m1}$ and $C_{a,m2}$ in the example shown in Fig. 4.9:

$$\begin{aligned} C_{a,m0}&: \quad 8 \\ C_{a,m1}&: \quad 11 \\ C_{a,m2}&: \quad 1292 \end{aligned}$$

At first sight this appears to heavily tamper the measurement result. However, this gap is implicitly closed when computing the TDOAs because the snapshots corresponding to the other transmission modes are likewise biased. This is shown as an example for the TDOA $t_{ma} - t_{mb}$. In transmission mode m, Eq. 4.15 becomes

$$S_{mq} = \sum_{j=1}^{N_B} \sum_{i=0}^{N_{m,j}-1} \left(i \cdot T_m + (3j+2)T_B + t_{mq} + t_{s0,m} + \Delta t_m - \Delta t_q + \Phi \right), \quad q \in \{a, b\}. \tag{4.22}$$

The term $3j \cdot T_B$ corresponds to the time gap between two mode m slots. The offset $2 \cdot T_B$ is caused by the sequence order displayed in Fig. 4.8. According to this order, modes b and c are entered before m, thus adding an offset of $2 \cdot T_B$.

When we apply the central limit theorem to Eq. 4.22 and divide by N_m, the expected value

[9] except once, before the entire measurement

of the resulting random variable $\hat{S}_{mq}$ becomes

$$E\left(\hat{S}_{mq}\right) = (3j+2)T_B + t_{mq} + t_{s0,m} + \Delta t_m - \Delta t_q + \sum_{j=1}^{N_B} \frac{(N_{m,j}-1)T_m + T_q}{2}.$$

When computing $E\left(\hat{S}_{ma}\right) - E\left(\hat{S}_{mb}\right)$, we obtain

$$\begin{aligned} E\left(\hat{S}_{ma}\right) - E\left(\hat{S}_{mb}\right) &= t_{ma} - t_{mb} - \Delta t_a + \Delta t_b + \frac{N_B}{2}(T_a - T_b), \quad \text{or} \\ t_{ma} - t_{mb} &= E\left(\hat{S}_{ma}\right) - E\left(\hat{S}_{mb}\right) + \Delta t_a - \Delta t_b - \frac{N_B}{2}(T_a - T_b). \end{aligned} \tag{4.23}$$

According to Eq. 4.19, we eliminate the clock offsets Δt_b and Δt_a by reference measurement c:

$$\begin{aligned} E\left(\hat{S}_{ca}\right) - E\left(\hat{S}_{cb}\right) &= t_{ca} - t_{cb} - \Delta t_a + \Delta t_b + \frac{N_B}{2}(T_a - T_b), \quad \text{or} \\ \Delta t_a - \Delta t_b &= E\left(\hat{S}_{cb}\right) - E\left(\hat{S}_{ca}\right) + t_{ca} - t_{cb} + \frac{N_B}{2}(T_a - T_b). \end{aligned} \tag{4.24}$$

Finally, by inserting Eq. 4.24 into Eq. 4.23, we obtain

$$t_{ma} - t_{mb} = E\left(\hat{S}_{ma}\right) - E\left(\hat{S}_{mb}\right) - E\left(\hat{S}_{ca}\right) + E\left(\hat{S}_{cb}\right) + t_{ca} - t_{cb}. \tag{4.25}$$

Comparing Eq. 4.25 with Eq. 4.18, we can see that all offsets and gaps introduced with the resynchronizing measurement procedure are implicitly eliminated. Hence, no modifications must be performed.

In the following section we will discuss how the system behaves if the number of mode switches is modified while the overall measurement time T_{sim} is constant. The more mode switches are accomplished, the less measurements will be performed in the overall simulation time because more time is reserved for switch delays. This context is depicted in Fig. 4.10. In an experiment, the measurement error is analyzed when the number of burst cycles is traded for the length of the bursts. The switch delay $t_{SD} = 50\,\mu s$ and the total measurement time $T_{sim} = 60\,ms$ were kept constant. All clocks have 12 ps RMS jitter. The possible range of burst numbers for this configuration is $N_B \in \{1, \ldots, 399\}$[10]. For the lower limit ($N_B = 1$), each transmission mode is entered only once. The upper limit ($N_B = 399$) is defined by the following constraint: Every slot b, c and m is entered N_B times, i.e. the overall measurement time is divided into $3 \cdot N_B$ time slots. Into each time slot, the switch delay must fit. If the switch delay is equal to or exceeds the burst time, then no pulses can be transmitted at all. Therefore, $T_{sim} > 3 \cdot N_B \cdot t_{SD}$ must hold, and $N_B = 399$ is the largest possible number.

To account for the random impact of clock jitter, each configuration was simulated 10 times. Accordingly, the ordinate of the plot shown in Fig. 4.10 is the standard deviation of measurement errors, $\sigma_{\Delta d_1}$ and $\sigma_{\Delta d_2}$.

It can be seen that the error decreases when the burst cycle length is decreased, because of less jitter accumulation. Interestingly, this is true even though for less burst cycles, i.e.

[10]The abscissa shows $N_B/2$ instead of N_B for reasons of the internal architecture of the simulator.

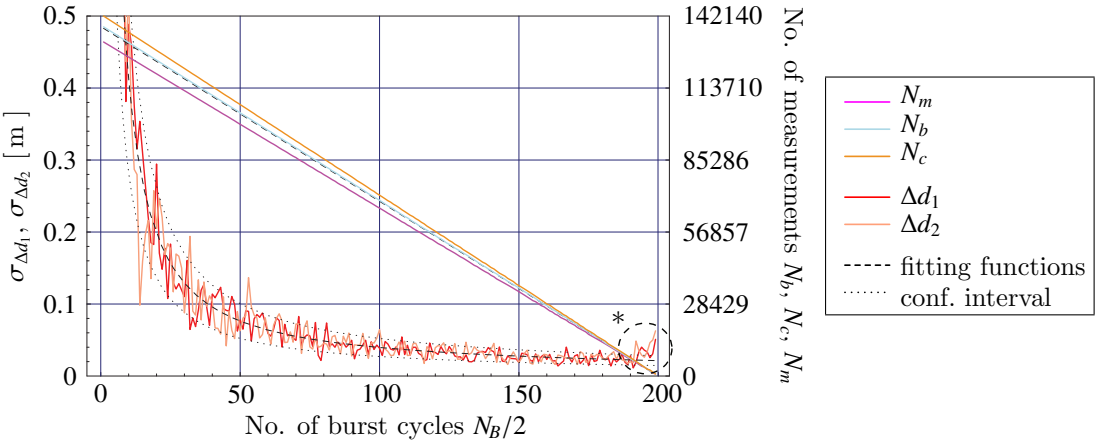


Figure 4.10: Measurement error vs. number of burst cycles, with constant overall measurement time $T_{sim} = 60\,\text{ms}$ and switch delay $t_{SD} = 50\,\mu\text{s}$

smaller N_B, a larger number of overall measurements N_b, N_c and N_m can be performed. In this device configuration, it is wiser to delimit jitter than to perform additional measurements. This result is particularly important when stringent energy conditions must be met. In this case, the number of pulse transmissions must be limited. Fortunately, as was shown above, this does not result in a loss of precision as long as clocks are resynchronized frequently.

It can also be observed that when the number of burst cycles N_B approaches its maximum, the measurement error increases. This behavior – indicated by a dashed circle (*) in Fig. 4.10 – can be explained in the following way: When N_B approaches its maximum, the burst time T_B approaches the switch delay, t_{SD}. Only in the remaining time, i.e. $T_B - t_{SD}$, pulses can be emitted. As we increase N_B, two trends work against each other: Fewer measurements can be performed, i.e. the result is degraded. On the other hand, less jitter is accumulated because the system is more frequently resynchronized and hence, the result is improved. From a certain point of N_B onwards, too few measurements can be performed to compensate the jitter and the low clock resolution. As a result, for very high N_B, the measurement precision decreases.

In the next set of simulations, the impact of different switch delays t_{SD} (10 µs and 50 µs) at different simulation times T_{sim} (60 ms and 6 ms), for various amounts of jitter, has been examined. The results are shown in Fig. 4.11. In these plots, a logarithmic scale was introduced in order to cover a larger range of burst cycles N_B. Three aspects are covered in these experiments:

1. In Fig. 4.11(a) and 4.11(b), the switch delay t_{SD} is 10 µs and 50 µs, respectively, while the overall measurement time $T_{sim} = 60\,\text{ms}$ is constant. As a result, more mode switches can be accomplished in case (a). It can be observed that despite the different switch

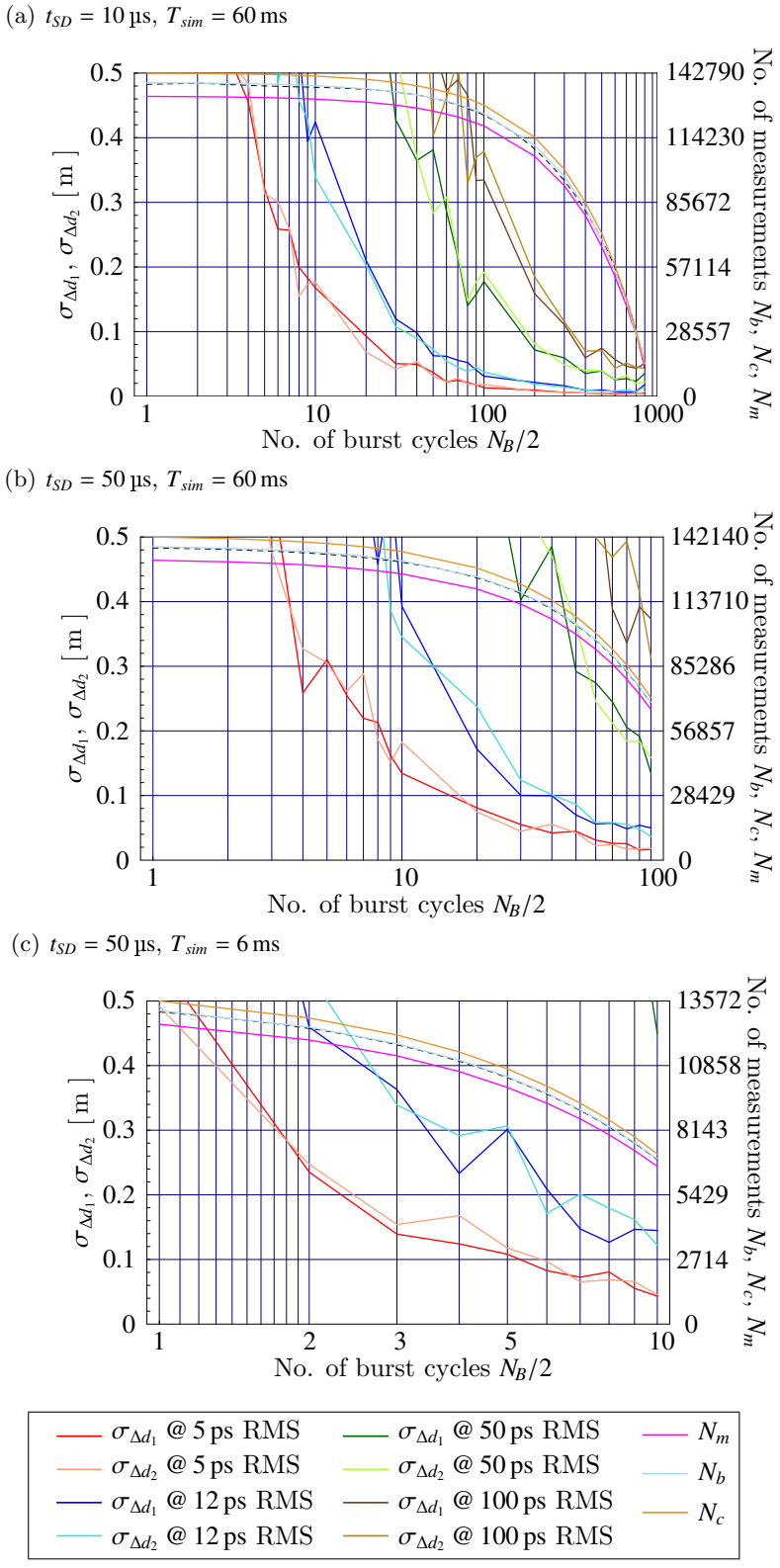


Figure 4.11: Measurement error vs. number of burst cycles for various device configurations and different amounts of jitter

delays, the resulting precision is very similar. As an example we can compare the errors at $N_B/2 = 100$: In both cases (a) and (b), we can read $\sigma_{\Delta d_1} \approx \sigma_{\Delta d_2} \approx 0.02$ for 5 ps RMS jitter, $\sigma_{\Delta d_1} \approx \sigma_{\Delta d_2} \approx 0.04 \ldots 0.05$ for 12 ps RMS jitter, etc. At the same time, in case (a) significantly more single measurements are performed (approx. 131,000 in (a) and 70,000 in (b), respectively). Yet, these additional measurements do not improve the result.

2. In Fig. 4.11(c) the measurement time is reduced to 6 ms. The switch delay is kept to 50 µs, as in Fig. 4.11(b). Accordingly, fewer mode switches can be performed in case (c). When comparing (b) and (c), an interesting observation can be made: With much less effort (i.e. number of measurements) almost the same precision can be achieved.

3. Furthermore, the general impact of clock jitter is shown in Fig. 4.11. As it could be expected, a higher jitter results in a lower precision of the estimate.

The following rule of thumb can be stated: The longer the overall measurement time, and the shorter the switch delay, the higher is the achievable precision. However, precision can be traded for measurement time and power. The most relevant requirement to achieve good results is to resynchronize the system as often as possible.

The plots in Fig. 4.11 show results for a *given measurement time*. When designing a localization system, this parameter is typically not available. Contrary to the measurement time, the switch delay – which is a physical constraint of the transceiver hardware – is usually given. Regarding this parameter as fix, we can now determine the maximum precision depending on the measurement time. In an example, a localization system was simulated at different measurement times T_{sim} ($3 \ldots 80$ ms in a logarithmic scale) and the minimum error acquired. The achievable minimum error was found in the following way: The average $\sigma_\Delta(N_B)$ of the errors d_1 and d_2 was determined for all values of N_B, with

$$\sigma_\Delta(N_B) = \frac{\sigma_{\Delta d_1}(N_B) + \sigma_{\Delta d_2}(N_B)}{2}.$$

A fitting function to this average was determined (see Fig. 4.10). The region of high N_B where the error increases again (dashed circle in Fig. 4.10) can be neglected because the fitting function is almost constant here, and hence no false minimum is generated. A confidence interval was then defined around the fitting curve, the width of which was chosen according to the number of simulations $N_S = 10$ that were run to compute statistics. Hence, the interval is limited by

$$\sigma_\Delta^\pm(N_B) = \sigma_\Delta(N_B) \left(1 \pm \frac{1}{\sqrt{N_S}}\right).$$

The upper limit of this interval, $\sigma_\Delta^+(N_B)$, was chosen as the guaranteed maximum measurement error for a particular number of burst cycles N_B. Finally, the minimum of this limiting curve,

$$\sigma_{\Delta,min} = \min\left(\sigma_\Delta^+(N_B)\right)$$

was determined. The entire procedure was repeated for T_{sim} ranging from 3 to 80 ms. In this example, a switch delay of 50 µs was chosen. The result is shown in Fig. 4.12. This plot

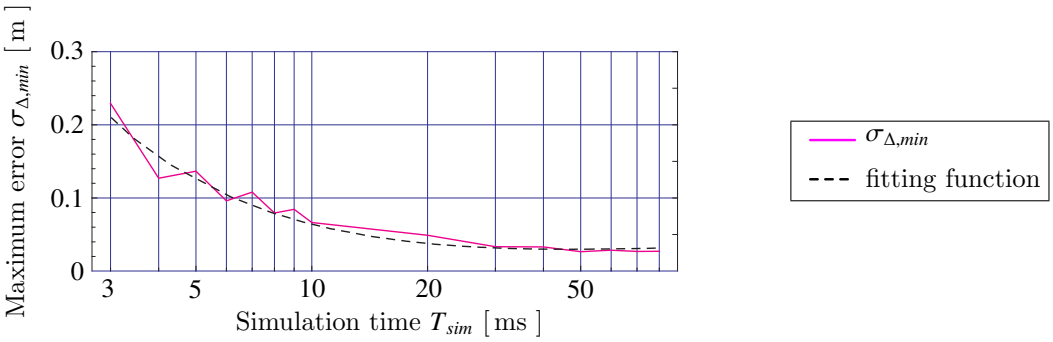


Figure 4.12: Measurement error vs. simulation time, $t_{SD} = 50\,\mu s$. All clocks have 12 ps RMS jitter

represents the maximum precision that can be achieved for a certain simulation time. Based on this curve, we can now determine the measurement time required to achieve the desired resolution.

4.5 Frequency bias compensation

Until now, only short term inaccuracies of the clocks have been considered. In this section, an approach to compensate long term frequency biases of all participating units is presented. We assume that during the time span of our overall measurement, the clocks do not drift, i.e. all clock frequencies are constant. This is a legal assumption since clock drift is a long term effect and the measurement takes less than a second. For the bias compensation approach, two modifications must be made:

- We introduce a fourth reference unit d, located at (x_d, y_d), which acts similar to the mobile unit m: Reference unit d emits a signal which is received by reference units a, b and c. Likewise, we perform TDOA measurements among a and b as well as a and c, respectively. The resulting distance differences are denoted $d_{d,ab}$ and $d_{d,ac}$. Since the position of d is known, the geometric arrangement can be exploited to compensate long term jitter.

  To account for the deviant clock frequencies, we introduce the clock speed correction factors c_a, c_b and c_c. Unit d need not be compensated because it only acts as a transmitter[11]. The mobile unit m need not be considered for frequency bias compensation at all for the same reason.

- The bias compensation measurement described above is performed a second time. By assuming that the measurements must yield the same results, the clock speed of all devices can implicitly be obtained.

[11] Remember that in hyperbolic localization, the emission time is non-relevant.

For simplicity, the computation of the correction factors is shown for clocks that are *ideal* but *biased in frequency*. Later on, the time averaging approach based on non-ideal, low frequency oscillators discussed in Sec. 4.3 is considered.

With c_a, c_b and c_c and in analogy to Eq. 4.1, 4.2, 4.11 and 4.12[12], we obtain the distance differences

$$d_{d,ab} = \left(c_a\, t_{r,da} - c_b\, t_{r,db} + t_{ac} - t_{bc} - c_a\, t_{r,ca} + c_b\, t_{r,cb}\right) \cdot v = l_{ad} - l_{bd} \quad \text{and} \tag{4.26}$$

$$d_{d,ac} = \left(c_a\, t_{r,da} - c_b\, t_{r,dc} + t_{ab} - t_{bc} - c_a\, t_{r,ba} + c_b\, t_{r,bc}\right) \cdot v = l_{ad} - l_{cd}. \tag{4.27}$$

As mentioned above, the measurement is repeated after some time. The variables referring to this second measurement are indicated by a hat:

$$\hat{d}_{d,ab} = \left(c_a\, \hat{t}_{r,da} - c_b\, \hat{t}_{r,db} + t_{ac} - t_{bc} - c_a\, \hat{t}_{r,ca} + c_b\, \hat{t}_{r,cb}\right) \cdot v = l_{ad} - l_{bd} \tag{4.28}$$

$$\hat{d}_{d,ac} = \left(c_a\, \hat{t}_{r,da} - c_b\, \hat{t}_{r,dc} + t_{ab} - t_{bc} - c_a\, \hat{t}_{r,ba} + c_b\, \hat{t}_{r,bc}\right) \cdot v = l_{ad} - l_{cd} \tag{4.29}$$

Thus, when equating Eq. 4.26 and 4.28, we get

$$c_a \left(t_{r,da} - \hat{t}_{r,da} - t_{r,ca} + \hat{t}_{r,ca}\right) + c_b \left(-t_{r,db} + \hat{t}_{r,db} + t_{r,cb} - \hat{t}_{r,cb}\right) = 0, \quad \text{or}$$

$$c_b = c_a \cdot \frac{t_{r,da} - \hat{t}_{r,da} - t_{r,ca} + \hat{t}_{r,ca}}{t_{r,db} - \hat{t}_{r,db} - t_{r,cb} + \hat{t}_{r,cb}} = c_a \cdot F_B, \tag{4.30}$$

where we introduce F_B for convenience:

$$F_B = c_a \cdot \frac{t_{r,da} - \hat{t}_{r,da} - t_{r,ca} + \hat{t}_{r,ca}}{t_{r,db} - \hat{t}_{r,db} - t_{r,cb} + \hat{t}_{r,cb}}$$

Furthermore, Eq. 4.26 can be written as

$$c_a \left(t_{r,da} - t_{r,ca}\right) + c_b \left(-t_{r,db} + t_{r,cb}\right) = \frac{l_{ad} - l_{bd}}{v} - t_{ac} + t_{bc}. \tag{4.31}$$

Finally, Eq. 4.30 can be inserted into Eq. 4.31. Solving for the correction factor c_a yields

$$c_a = \frac{\frac{l_{ad} - l_{bd}}{v} - t_{ac} + t_{bc}}{t_{r,da} - t_{r,ca} - \left(t_{r,db} - t_{r,cb}\right) \cdot F_B}. \tag{4.32}$$

c_b can be computed from Eq. 4.30. Likewise, c_c can be expressed as a function of c_a when equating Eq. 4.27 and 4.29.

In these considerations, it was assumed that clocks are ideal (but biased from their nominal frequencies). The bias compensation scheme depicted above can easily be extended for non-ideal, low frequency clocks. Two steps must be undertaken for this purpose:

[12] With these equations, the distance differences d_1 and d_2 can be obtained when a signal is emitted from the mobile unit m. Here, m is substituted by d. As a result, the distance differences $d_{d,ab}$ and $d_{d,ac}$ are computed.

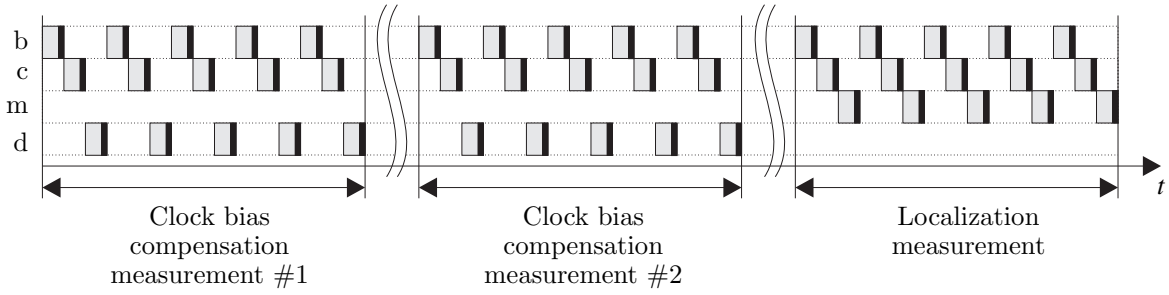


Figure 4.13: Transmission pattern accounting for long term clock frequency bias and jitter

- First, we introduce the *nominal* clock period $T_{p,0}$ for unit p. The *actual* clock period is denoted T_p. The nominal and the actual value are linked by the clock correction factor c_p:

  $$T_p = T_{p,0} \cdot c_p, \quad p \in \{a, b, c\}$$

  The equations introduced in Sec. 4.3 remain unchanged.

- The correction factors c_a, c_b and c_c are obtained by the two clock bias compensation measurements. These measurements must be inserted into the transmission pattern. One possible arrangement is shown in Fig. 4.13. Prior to any localization measurement, two clock bias compensation measurements are performed. These measurements are built up in the same way as any other measurement. I.e. smaller bursts are performed in order to restrict jitter accumulation. Finally, one or multiple localization measurements can be accomplished.

  It is suggested to frequently repeat the clock bias measurements, e.g. once a minute. Thus, temperature variations (which typically happen at a very slow rate), can be compensated.

With the transmission pattern shown in Fig. 4.13, all aspects discussed in the previous sections are considered:

- Inexpensive low frequency clocks are used for time measurement.
- Clocks are implicitly synchronized.
- Clock jitter is delimited.
- Long term oscillator drift is compensated.

In this section, a novel approach on time measurement in a hyperbolic localization system using low accuracy and low cost oscillators was introduced. The basic idea of the method is to transmit contiguous trains of pulses for TDOA measurements, instead of a single pulse.

The method has two major advantages:

- A high location accuracy is obtained in spite of using relatively low frequency and unstable or drifting clock oscillators. The robustness against clock jitter is achieved by transmitting short bursts of pulses following a certain transmission pattern. As a consequence, localization equipment can be implemented in inexpensive standard CMOS without the need for sophisticated high speed circuitry.
- The participating units' clocks are implicitly synchronized which further reduces system complexity and cost. Simulation results have been provided that indicate how the achieved precision depends on various factors such as burst length and jitter.

4.6 Error estimation

In the real world, TDOA measurements are subject to noise. Until now, only the error of the distance differences, Δd_1 and Δd_2 have been considered. However, in order to judge a system, the *resulting localization error* is a more feasible measure. Therefore, it is investigated how the computed position (x_m, y_m) of the mobile unit varies for assumed maximum measuring errors of Δd_1 and Δd_2, with $\Delta d_1 = \Delta d_2 = \Delta d$. This position shift was examined in two different ways, the first of which is a numerical and the second one an analytical solution based on Gaussian error propagation. In the following, these two approaches will be described and compared. Subsequently, accuracy distributions in the (x, y) plane are presented that show how reference units can be deployed in advantageous and disadvantageous settings.

4.6.1 Numerical error propagation estimation

For the numerical estimation, the following method is applied: For all points (x, y) in the plot range, distance differences d_1 and d_2 are computed according to Eq. 4.1 and 4.2, respectively. These values represent the *exact* distance differences. In order to account for measurement errors, the maximum error Δd is applied (see Fig. 4.14): In a loop, different error values Δd_1 and Δd_2 ($\Delta d_1, \Delta d_2 \in \{-\Delta d, \Delta d\}$) are added, resulting in four pairs of erroneous distance differences $\left(d_1^{(i)}, d_2^{(i)}\right)$, with $i \in \{1, 2, 3, 4\}$. These pairs represent the worst case of measurements. For each pair, the corresponding position $\left(x_m^{(i)}, y_m^{(i)}\right)$ is calculated[13]. This position is compared to the original position (x_m, y_m) by determining the Euclidean distance

$$\Delta e(x_m, y_m, i) = \sqrt{\left(x_m - x_m^{(i)}\right)^2 + \left(y_m - y_m^{(i)}\right)^2}.$$

Finally the biggest deviation $\Delta e_{max}(x_m, y_m)$ is chosen. This result represents the largest possible deviation from the true position. The procedure is repeated until all points (x, y) (at certain intervals) have been covered.

[13]The direct calculation method described in App. B was used.

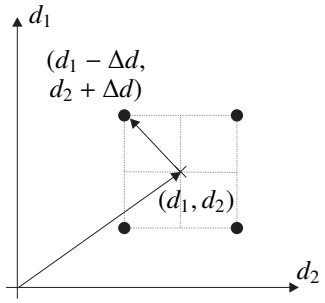


Figure 4.14: Numerical approach

Fig. 4.15(a) shows a plot based on this algorithm. The reference units are located at $(-3.0\,\mathrm{m}, 0\,\mathrm{m})$, $(4.5\,\mathrm{m}, 2.5\,\mathrm{m})$ and $(-1.0\,\mathrm{m}, 8.0\,\mathrm{m})$. The assumed maximum error Δd is 0.1 m.

The graph has the shape of a large flat bowl with a minimum of approx. 0.1 m. At positions close to the reference units, the graph rises steeply, indicating that the computed position (x_m, y_m) is highly susceptible to inaccuracies of the distance difference measurements.

4.6.2 Analytical error propagation estimation

In a second approach, an analytical description of the position error estimation was elaborated. For this purpose we introduce two functions f^x and f^y, where

$$f^x(d_1, d_2) = x_m, \quad f^y(d_1, d_2) = y_m. \tag{4.33}$$

For a pair of measured (erroneous) values D_1 and D_2, the deviation $(\Delta f^x, \Delta f^y) = (\Delta x_m, \Delta y_m)$ from the *true* values $f^x(d_1, d_2)$ resp. $f^y(d_1, d_2)$ can be determined. For the analytical approximation[14], we consider the linear part of the Taylor series of f^x and f^y and determine the absolute deviation[15]:

$$|\Delta x| = \left|\Delta f^x\right| \approx \left|\frac{\partial f^x}{\partial d_1}\Delta d_1 + \frac{\partial f^x}{\partial d_2}\Delta d_2\right|$$

Note that equally signed errors Δd_1 and Δd_2 do not necessarily shift the outcome f^x in the same direction. In certain cases, they might expunge each others' influences. As we want to determine an upper bound estimate of the resulting localization error, the *absolute* values of the partial derivatives are considered. Furthermore, we again assume that both measurements D_1 and D_2 are equally deviated, i.e. $\Delta d_1 = \Delta d_2 = \Delta d$:

$$\left|\Delta f^x\right| \lesssim \left(\left|\frac{\partial f^x}{\partial d_1}\right| + \left|\frac{\partial f^x}{\partial d_2}\right|\right) \cdot |\Delta d| \tag{4.34}$$

[14] according to Gaussian error propagation

[15] shown for the x-direction.

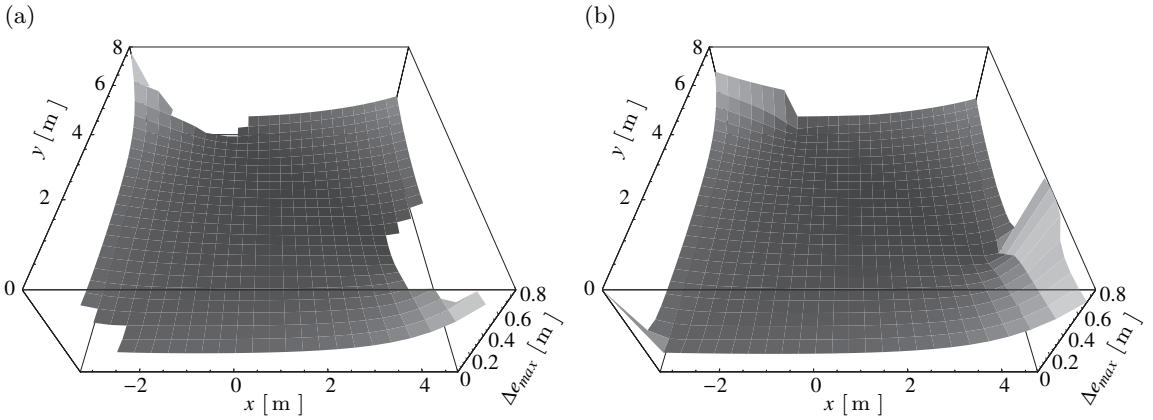


Figure 4.15: Localization error distribution for a measurement error $\Delta d = 0.1\,\mathrm{m}$ as a result of the numerical (a) and analytical (b) approach. The plot range of (a) includes three areas where the position (x_m, y_m) cannot be unambiguously computed. Hence these areas are excluded in the estimation (see also Fig. 4.2). The analytical estimation in (b) does not consider these areas.

A respective approximation as shown for the x-direction can be set up for $\Delta f^y = \Delta y$.

To determine the position (x_m, y_m), we use the direct calculation method according to App. B. Unfortunately, this solution has a high complexity. As a result, determining the partial derivatives of f^x and f^y is a difficult task. As an alternative, these derivatives can easily be found by determining the derivative of the inverse function of

$$\vec{f}(d_1, d_2) \;=\; \begin{pmatrix} f^x(d_1, d_2) \\ f^y(d_1, d_2) \end{pmatrix} \;=\; \begin{pmatrix} x_m \\ y_m \end{pmatrix}.$$

For this purpose, we use the following rule:

$$\vec{f}'(d_1, d_2) \;=\; \left(\left(\vec{f}^{-1} \right)' (x_m, y_m) \right)^{-1} \tag{4.35}$$

The inverse function of $\vec{f}(d_1, d_2)$ is given by

$$\vec{f}^{-1}(x_m, y_m) \;=\; \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}. \tag{4.36}$$

The derivative of this inverse function can easily be determined because d_1 and d_2 have a very low complexity:

$$\begin{aligned} d_1 &= l_{am} - l_{bm} \\ &= \sqrt{(x_m - x_a)^2 + (y_m - y_a)^2} - \sqrt{(x_m - x_b)^2 + (y_m - y_b)^2} \\ d_2 &= l_{am} - l_{cm} \\ &= \sqrt{(x_m - x_a)^2 + (y_m - y_a)^2} - \sqrt{(x_m - x_c)^2 + (y_m - y_c)^2} \end{aligned}$$

Therefore, using Eq. 4.35 and 4.36, the derivative of $\vec{f}(d_1, d_2)$ can be expressed as

$$\vec{f}'(d_1, d_2) = \begin{pmatrix} \frac{\partial d_1}{\partial x_m} & \frac{\partial d_1}{\partial y_m} \\ \frac{\partial d_2}{\partial x_m} & \frac{\partial d_2}{\partial y_m} \end{pmatrix}^{-1}$$

The partial derivatives $\partial f^x/\partial d_1$ and $\partial f^x/\partial d_2$ can be obtained from the elements of the inverted matrix:

$$\vec{f}'(d_1, d_2) = \begin{pmatrix} \frac{\partial f^x}{\partial d_1} & \frac{\partial f^x}{\partial d_2} \\ \frac{\partial f^y}{\partial d_1} & \frac{\partial f^y}{\partial d_2} \end{pmatrix}$$

When inserting $\partial f^x/\partial d_1$ and $\partial f^x/\partial d_2$ into Eq. 4.34, we obtain an approximation of the maximum positioning error in x-direction, denoted Δx_m. In the same way, the partial derivatives $\partial f^y/\partial d_1$ and $\partial f^y/\partial d_2$, and thus, the maximum error in y-direction, Δy_m, can be determined.

In a final step we compute the Euclidean distance between the true point (x_m, y_m) and the maximum erroneous point $(x_m + \Delta x, y_m + \Delta y)$:

$$\Delta e_{max}(x_m, y_m) = \sqrt{{\Delta x_m}^2 + {\Delta y_m}^2}$$

$\Delta e_{max}(x_m, y_m)$ can be plotted in an x-y-diagram, as shown in Fig. 4.15(b).

To compare the numerical with the analytical approach we exclude the ambiguous areas around the reference units. The analytical approximation is very close to the numerical results: In the range plotted in the example we observe a difference of less than ±10% for most of the points. The analytical approach is a slim and efficient closed-form solution to compute the maximum location error distribution of a TDOA based localization system. It can be used to analyze or design the geometrical setup such that all precision constraints demanded by an application are met.

4.6.3 System setup strategies

Based on the analytical approach described in the previous section, it is demonstrated how the error propagation analysis may help to design the localization system, i.e. how to choose positions for the reference units. Fig. 4.16 shows the error propagation plots of three different setups of reference units in a rectangular room (indicated by a dashed line). Reference units a (lower left corner) and c (upper left corner) are fixed and b is varied. Again, dark areas correspond to positions highly robust against sensor noise while in the bright areas the system is susceptible to errors. Therefore, in the bottom right corner of the leftmost plot even small deviations of d_1 and d_2 can lead to enormous position inaccuracies: a maximum error

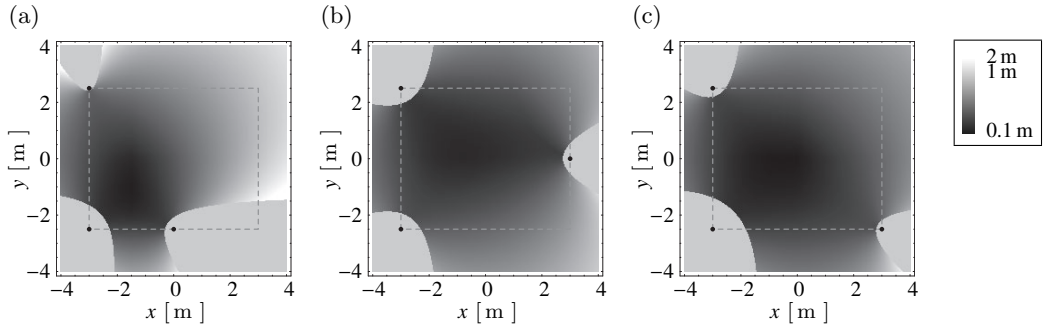


Figure 4.16: Error propagation plots for different reference unit arrangements. For the maximum error, $\Delta d = 0.1\,\mathrm{m}$ was assumed. Please note that in order to achieve a large plot range, $\log(\Delta e_{max}(x_m, y_m))$ was used for these plots, unlike those in Fig. 4.15.

$\Delta d = 0.1\,\mathrm{m}$ results in a maximum location deviation of approx. $1.0\,\mathrm{m}$, i.e. 1000% of the original error. In both other configurations, (b) and (c), a much higher robustness can be observed at this location.

This example shows how a room can be 'illuminated' for different purposes. Maybe for one application it is reasonable to achieve the highest precision in the lower left corner. Maybe there is furniture in the upper right corner, so localization is rarely performed there. The most homogeneous result can be achieved when the reference units are located as in the middle and right-hand plot. Please note that – depending on the application – a homogeneously illuminated room need not necessarily be desired. In general, the reference units should be placed at the borders of the room in order to minimize the ambiguous and highly error susceptible areas.

Chapter 5

AOA measurement using an IR detector array

A well known alternative to hyperbolic localization, discussed in Ch. 4, is triangulation. This method is based on the measurement of angles of incidence of signals transmitted between the mobile unit and fixed reference units. In Sec. 2.4.7, different methods of measuring the angle of arrival (AOA) of a signal against an arbitrary base line have been presented. Using infrared technology, the angle of arrival can be determined based on either time difference of arrival measurements or received signal strength. In this section, a new all-static method for angle of arrival measurement based on signal strength is introduced. The approach exploits the directivity of an array of infrared sensors that are positioned in a ring-shaped arrangement. To detect the angle of arrival, we sweep through the field of vision of the detector array and try to find the angle yielding the maximum signal strength. However, the sensors are not *physically* rotated but by means of discrete signal processing. For this purpose, each sensor reading is weighted by a certain factor for each angle of the sweep. Due to the non-uniform radiant sensitivity of the sensors and the angular displacement against each other, an incident ray of radiation has a different impact on the reading of each sensor. Therefore, with properly chosen sensor weights combined with signal processing techniques, it is possible to approximate the original radiation pattern received by the detector. Finally, the angle yielding the maximum signal strength is extracted from this reconstructed pattern. The method presented here is all-static. Only a one time snapshot of sensor readings, i.e. one value per sensor, must be taken. All processing, including the angular sweep, is performed offline, subsequently. The approach is suited for multiple sensor technologies. The most common and inexpensive IR detectors are photodiodes. However, the method is not limited to these devices. Thermal detectors such as thermopiles have advantages over photodiodes at wavelengths of several µm.

Name(s)	*Wavelength range*
IR-A, near IR, NIR	780 ... 1400 nm
IR-B	1.4 ... 3 µm
IR-C, far IR	3 µm ... 1 mm

Table 5.1: CIE vocabulary for spectral regions

Terminology for signals in the angular domain

A remark must be made concerning the terminology used in this chapter. In common applications of digital signal processing, signals are described by a series of numbers, indexed by n, e.g. $b(n)$. n often represents discrete time stamps $n \cdot T$, where T represents the sampling interval. Therefore $b(n)$ describes a signal whose value changes in *time*. Its frequency spectrum $B(k)$ can be determined using the discrete Fourier transform.

Besides time, there exist other domains where discrete signals are accordingly described. E.g., in image processing, $b(n_x, n_y)$ usually corresponds to the luminance of pixels in digital images. In this case, n_x and n_y represent spatial distances in the image plane. In image processing, the term *spatial frequency* has been introduced for the frequency domain.

In this chapter, we consider signals in the *angular* domain. I.e., n represents discrete angular steps $n \cdot \Delta r$, with Δr being the angular sampling period. For the discrete Fourier transform $B(k)$ of a signal $b(n)$, we introduce the term *angular frequency spectrum*. Please note that this denotation is not related to the term *angular frequency* commonly used in physics, mechanics and electrical engineering to describe various kinds of oscillation and rotation.

Infrared technology

Infrared technology is used in uncounted applications today. As early as in 1969, Hudson presented the infrared applications matrix, showing that various IR technologies can be applied to military, industrial, medical and scientific problems [46]. Back then, the major activity has been in military applications. Commercialization was long time limited or lagged due to the high cost of infrared technology. Thanks to the significant advances in detector devices and supporting electronics, major reductions of cost of infrared systems have been achieved. Today, the breadth of applications as laid out by Hudson, can be realized at reasonable cost. One such application is location estimation.

The term *infrared* refers to the range of wavelengths in the electromagnetic spectrum that lies between the visible spectrum and the microwaves. There is no unique definition of the boundaries of this range. Therefore, different classifications and labels for certain portions of the infrared range exist. The broadly accepted terminology standardized by the International Commission on Illumination (Commission Internationale de l'Eclairage, CIE) is shown in Tab. 5.1 [22].

An excellent book on radiometry covering the infrared spectrum was written by McCluney [61]. Infrared technology has also found broad coverage in indoor wireless data communication in the last 15 years. Good surveys of this field were written by Barry [9], Kahn [52], Street [83], Carruthers [19] and Singh et al. [79]. Two standards must be mentioned here: IrDA enables directed point-to-point data transfers of up to 16 Mbps at a range of 1 m [90]. It is aimed at wireless data transfer between mobile personal appliances like PDAs. Thanks to the low implementation cost it became the predominant wireless short range communication interface. However, IrDA links require an uninterrupted line of sight and devices must be pointed at each other. Due to this inconvenience, the acceptance of IrDA has never been great among users. The second standard, IEEE 802.11, presents the physical (PHY) and medium access (MAC) protocol layers for a diffuse infrared wireless LAN [1]. Due to the popularity of radio wireless LANs, the existence of infrared wireless networking is hardly perceived. The knowledge on infrared *communication* is a valuable source for adapting infrared to indoor *localization*.

The following sections are structured as follows: As a starting point, a 2D triangulation scheme based on AOA measurements is introduced (Sec. 5.1). Receiver and transmitter designs are briefly discussed in Sec. 5.2.1 and 5.2.2. The new AOA measurement approach is described in detail in Sec. 5.6.

5.1 Triangulation scheme

In Sec. 2.4.7, triangulation was introduced with the reference units acting as beacons. According to this approach, these reference units emit beacon signals which are received by one or multiple mobile units. Since the orientation of these mobile units is unknown beforehand, three reference units are needed. Alternatively, the mobile units can act as emitters, and the reference units determine the angle of arrival of a beacon signal emitted from these mobile units. In this case, the orientation of the receiving units, i.e., the reference units, is fixed. As a consequence only two reference units are required. This setup is illustrated in Fig. 5.1. In this arrangement, two reference units a and b are mounted at the walls of a room, at locations (x_a, y_a) and (x_b, y_b). Their field of view covers 180°, and their optical centerline (0°) is perpendicular to the wall they are mounted on. The mobile unit m emits a signal into all directions of the x-y-plane, which arrives at angles α_a and α_b. From the illustration, we can derive

$$\frac{x_m - x_a}{y_m - y_a} = \tan\alpha_a \quad \text{and} \quad \frac{x_b - x_m}{y_b - y_m} = \tan\alpha_b.$$

With these equations, the location of the mobile unit (x_m, y_m) can be obtained:

$$\begin{aligned} y_m &= \frac{x_a - x_b - y_a \tan\alpha_a + y_b \tan\alpha_b}{\tan\alpha_b - \tan\alpha_a} \\ x_m &= x_a + (y_m - y_a)\tan\alpha_a \end{aligned}$$

Note that this simple setting yields no solution if m resides on the direct line between a and b, since then $\tan\alpha_b - \tan\alpha_a = 0$ holds. We can avoid this situation by adding an additional

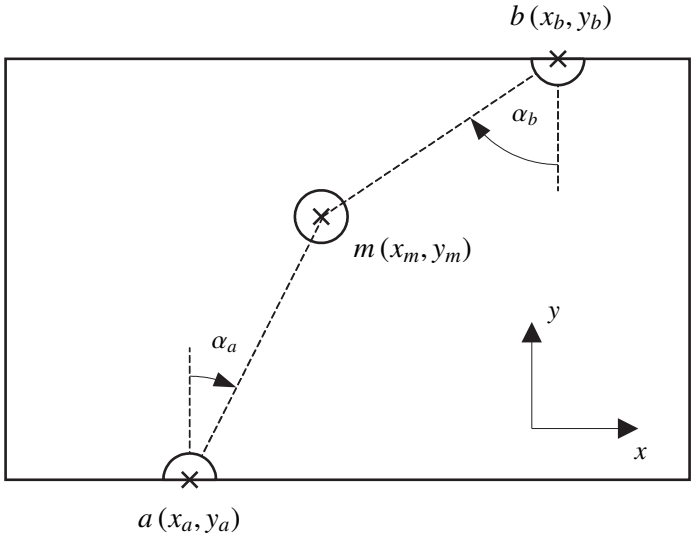


Figure 5.1: Geometrical arrangement for triangulation in the plane

reference unit. Alternatively, in case the equation cannot be solved, a Bayesian filter can predict the current position from the recent movement.

When using IR radiation for localization, two further prerequisites must be met:

- The line of sight between the mobile unit and both reference units must be uninterrupted.
- It must be assured that the mobile unit emits the signal into the horizontal plane.

The latter point can easily be achieved if the transmitter is mounted on top of a vehicle moving on wheels on a flat floor (e.g., an AGV or robot). It is much more complicated to implement such a transmitter in an active badge worn by a human. In this case, the emitter may incline according to the movement of the human body. Furthermore, a large angular range is blocked by the body the badge is attached to. Hence the transmitter is shadowed in these directions. Due to these limitations, active localization in the way presented here is aimed at vehicle and robot navigation. Alternatively, the human body may act as a natural emitter of infrared radiation. This scenario requires a different type of sensors, but the principle remains unchanged. An outlook concerning the localization of humans is given in Sec. 6.4.

5.2 Receiver and transmitter design

Against the background of infrared wireless LANs, a number of modifications were proposed for transmitter and receiver devices. The motivation was to achieve performance improvements in communication. Some of the proposed devices can also be used in a positioning

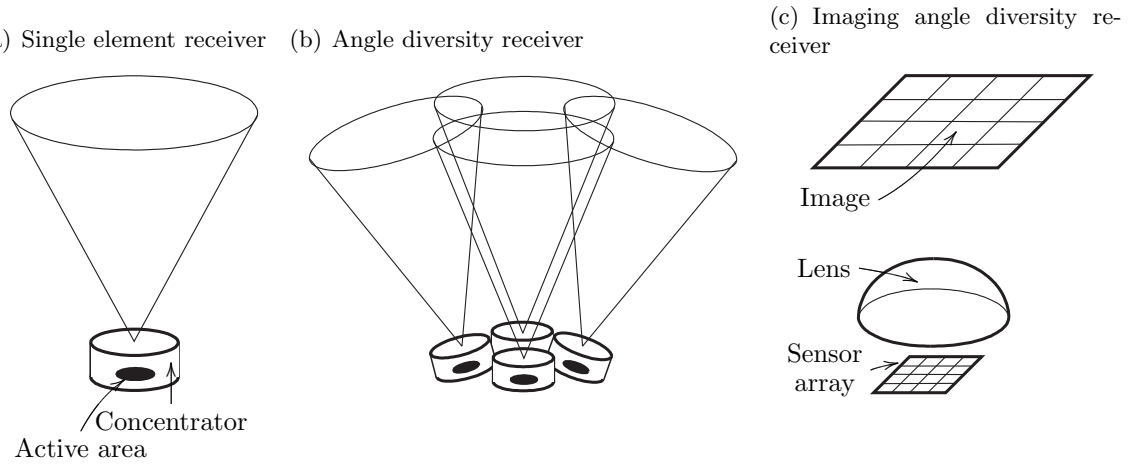


Figure 5.2: Infrared receivers

system. Among these are multi-beam transmitters and angle diversity receivers [84]. These devices are briefly introduced in Sec. 5.2.1 and 5.2.2.

5.2.1 Receiver design

In the following sections, we assume photodetectors as sensitive elements. These devices are inexpensive and can be operated uncooled, at least at wavelengths from 800 ... 1300 nm. Fig. 5.2 shows different kinds of IR receivers. *Single-element receivers* consist of an optical concentrator (a lens) which is coupled to the detector (Fig. 5.2(a)). In this receiver, besides the desired signal, delayed multipath components and ambient light are combined in a single electrical signal. An *angle diversity receiver* utilizes multiple detector elements that point into different directions (Fig. 5.2(b)). Depending on the directionality of the concentrators attached to the sensor elements, the output signals are coupled against each other to a certain degree. These receivers can significantly reduce the implications of ambient lighting and multipath dispersion, in case these unwanted signals are received from different directions than the expected signal. Unfortunately, each sensing element must be attached to its own optical concentrator and mounted in a precise angle such that it points in the desired direction. This effort makes these receivers unnecessarily clumsy and costly. These disadvantages are addressed by the *imaging angle diversity receiver*, first proposed by Yun and Kavehrad [96] (Fig. 5.2(c)). It consists of a single optical concentrator coupled to a planar array of sensing elements. This yields a sensor array which is both inexpensive and relatively straightforward to manufacture. Kahn et al. developed an experimental imaging receiver design for 100 Mbps indoor wireless optical communication [54]. It consists of a silicon pin photodiode array of 37 elements, three imaging lenses and an optical interference filter. With the arrangement

shown in Fig. 5.3, a half angle field of view of 45° is obtained. Unfortunately, at an image plane of 2.3 cm diameter, the scale of the device is rather large and hence very expensive.

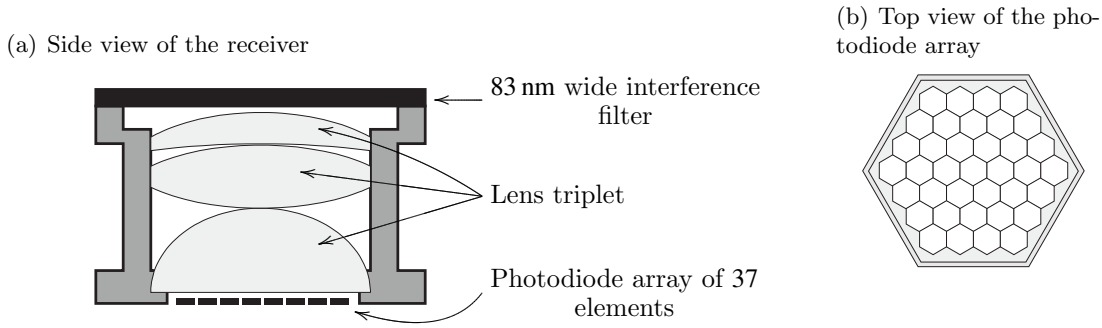


Figure 5.3: Experimental imaging receiver. Each pixel has hexagonal shape in order to employ chip area in the most expedient way.

5.2.2 Transmitter design

In the wavelength range of 880 ... 1300 µm, there exist two types of integrated transmitter devices, the light emitting diode (LED) and the laser diode (LD). Inexpensive variants of both of these devices are available.

LEDs are based on the spontaneous emission of radiation when injected charge carriers recombine in a forward biased pn-junction. The wavelength of this radiation depends on the band gap or impurities of the semiconductor structure. The spectral range of emission is limited, but not so much to be considered monochromatic. Thus an LED is an incoherent source. Typically, materials on the basis of GaAs are used in infrared LEDs. The spontaneous emission generates photons that travel in random directions, so the emission is isotropic. Therefore, only little of the theoretical output is available for use [91]. Infrared LEDs are also referred to as IREDs (Infrared emitting diode).

Laser diodes are semiconductor devices predominantly based on stimulated emission. A photon that travels across the highly degenerated pn-junction stimulates electron hole pairs to recombine and thus emit further photons of the same direction, polarization and phase as the first photon. A Fabry-Perot-Resonator (a set of two parallel mirrors) makes the laser act as an oscillator. Therefore, photons traverse the pn-junction many times before escaping the device. Light amplification in a laser diode enables a much higher radiation intensity than an LED. Furthermore, almost monochromatic and coherent radiation is generated. The laser mode is entered if the forward bias current exceeds the certain threshold.

One or more lenses, or a diffuser, can be attached to the radiation source in order to direct the radiation in one way or another. For concerns of eye safety, diffuse transmitters enable higher transmit powers than beam transmitters because only a small fraction of the emitted radiation is received by the human eye. Laser diodes have several advantages over

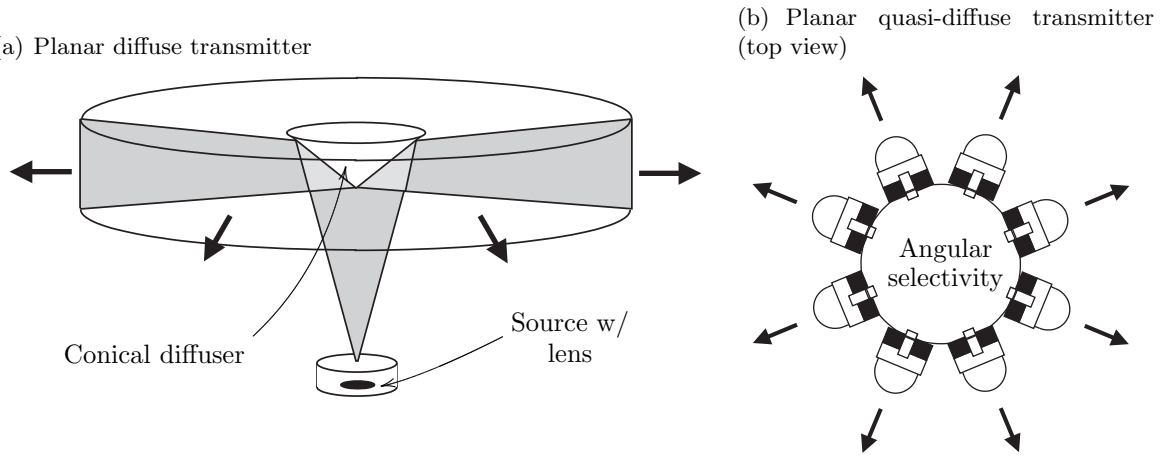


Figure 5.4: Planar transmitters mounted on top of the mobile unit

LEDs: they can emit at much higher optical power levels, and they usually yield better efficiency. Since the emitted light is almost monochromatic, they enable much narrower optical filters at the receiver, and thus reduce ambient noise. Unfortunately, they are usually much more expensive. Furthermore, laser operation is only achieved if the forward bias current threshold is exceeded. Therefore, the emitted optical power cannot be adjusted arbitrarily. However, this aspect is relevant mainly for communication links, when sub-carrier modulation techniques are considered.

In order to localize a mobile unit, e.g. a robot or AGV, it must emit a signal in all directions of the horizontal plane. The most simple and inexpensive design of this emitter is shown in Fig. 5.4(a). The collimated beam of a vertically mounted IR source is deflected into the horizontal plane by a conical diffuser. This diffuser must have good reflective properties, as is the case with e.g. smooth metallic surfaces. Because the radiation is scattered, the device can be operated at a fairly high transmit power level. However, for concerns of battery lifetime it is desirable to reduce the emitted power to an acceptable limit.

If the approximate direction of the closest AOA detectors is known, it may be beneficial to emit radiation only into these directions. This idea of angular selectivity can be realized with a planar multi-beam, quasi-diffuse emitter, as shown in Fig. 5.4(b). In this device, multiple IR sources are circularly mounted in the horizontal plane. While this arrangement is more expensive than the diffuse planar emitter discussed above[1], it may save energy by driving only those sources relevant for localization.

[1]given that the same type of IR source (IRED or LD) is used

5.3 Infrared indoor channel

The infrared indoor channel has vastly been examined in various experiments. Various models for directed and diffuse communication links have been developed [10, 53, 64, 15, 20]. The performance of wireless indoor infrared systems is constrained in many ways: radiation propagation is characterized by high path loss and multipath dispersion. Often, there exists significant ambient natural (sunlight) and artificial (incandescent and fluorescent) lighting [65]. Both types of this background light induce shot noise in the receiving device. Furthermore, artificial lighting produces severe interference due to the optical power fluctuations impinging the photodiode. These variations in time are usually fast and large. Natural ambient light does not yield this effect since it is almost stationary, considering the relatively slow variations of intensity and solar movement. Interference can prevent infrared systems from operating at all or demand excessive transmission power levels. A large part of these perturbations can be eliminated by optical filtering. Furthermore, high pass electrical filtering is often performed since most of the interference is situated at low frequencies of the power spectrum. For example, electronic ballasts of fluorescent lamps yield strong signal components at 100 kHz (and – to a declining degree – its harmonics [65]).

The background radiation that passes the filter adds to the receiver's shot noise. Shot noise is the dominant type of receiver noise in the photo-conductive mode. It can be modeled as white, Gaussian and independent of the input signal $X(t)$ [11]. In addition to ambient noise, the amplifier noise of the receiver must be taken into account which is also signal independent and Gaussian. Therefore, the receiver noise $N(t)$ can be modeled Gaussian and signal independent. The indoor channel can be modeled as a baseband linear system, with input $X(t)$, output $Y(t)$, receiver responsivity R and an impulse response $h(t)$.

$$Y(t) = RX(t) * h(t) + N(t)$$

The signal to noise ratio (SNR) of a direct line of sight receiver is proportional to the square of the received optical power. As a consequence, IR links operate at high transmit power levels and a relatively limited range. The maximum transmitter power is constrained by concerns of eye safety and power consumption.

5.4 Angular sensor characteristic

The angular directional radiant characteristic of a photodiode strongly depends on the shape and optical properties of the lens attached to it. Usually, for reasons of less intricate manufacturing, this shape has uniaxial symmetry. The characteristic is commonly described by powers of the cosine function. Fig. 5.5(a) shows the angular sensitivity profile of a receiver. An alternative representation that is often used in documentation and datasheets is shown in Fig. 5.5(b). The sensitivity profile is defined by

$$d(\phi) = \begin{cases} \cos^{\kappa}\left(\phi \cdot \frac{90^\circ}{\tau} \cdot \frac{\pi}{180^\circ}\right), & \text{if } -\tau \le \phi \le \tau, \\ 0 & \text{otherwise,} \end{cases} \tag{5.1}$$

where κ is a factor indicating the directionality of the receiver. τ is measured in degrees and indicates the half angle of the field of view.

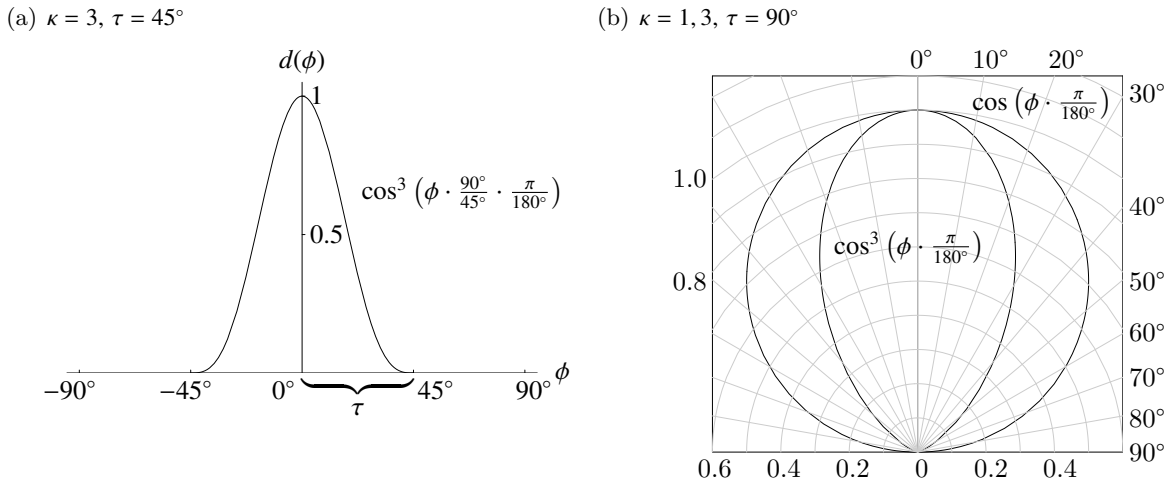


Figure 5.5: Relative radiant sensitivity $d(\phi)$ vs. angular displacement ϕ

5.5 Basic principle of scene sectoring

The goal of the system described here is to detect the angle of arrival of an IR beam by a detector array composed of a number of specially arranged sensors. In the approach, the directional characteristic of the single sensors is exploited. The incident radiation profile is reconstructed. A peak in the reconstructed signal indicates a source of radiation.

First, we consider a 360° field of view (FOV) variant of the detector, with L sensors rotated by a constant amount Φ. An example of this sensor arrangement for $L = 6$ sensors is shown in Fig. 5.6. The orientation of the sensors is defined by the vector $\boldsymbol{\rho} = \{\rho_l\}^{\mathrm{T}}$, where

$$\rho_l = -180° + l \cdot \Phi, \quad 0 \leq l < L. \tag{5.2}$$

For convenience, a shifted sensitivity profile function of each sensor, $d_l(\phi)$, is introduced as

$$d_l(\phi) = d(\phi - \rho_l). \tag{5.3}$$

Each sensor of the circular sensor arrangement converts the incident radiant flux to an electric current or voltage. Here, we introduce a proportional and dimensionless equivalent for the sensor readings, $x(\phi)$. From these sensor readings, we attempt to reconstruct the incident radiation profile $b(\phi)$. For now, we assume that only one source of radiation exists, as shown in the example in Fig. 5.7

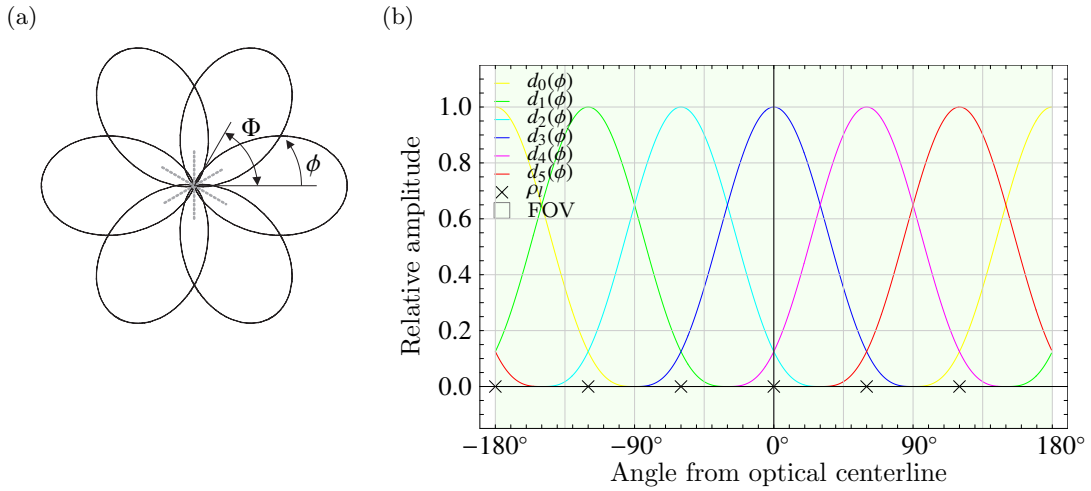


Figure 5.6: Circular sensor arrangement and corresponding sensitivity profiles ($L = 6$ sensors, sensor rotation $\Phi = 60°$). The green section indicates the field of view (FOV). Here, the FOV is the entire angular range, i.e. 360°.

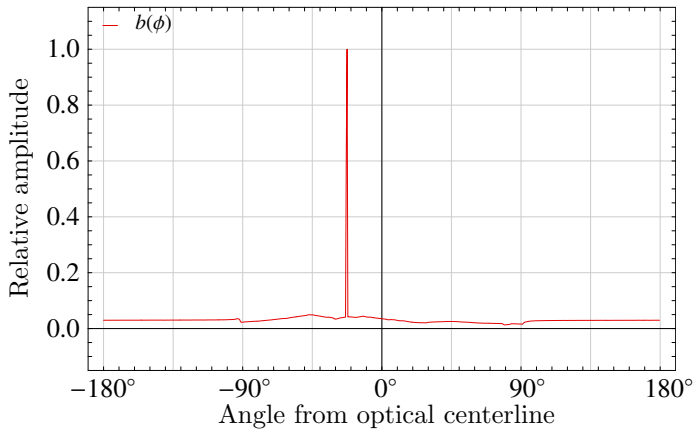


Figure 5.7: Incident radiation profile $b(\phi)$

The sensor arrangement depicted above can be mathematically described as a sampling system, with a sampling period of Φ. The signal flow of the device in the continuous signal space is shown on the left hand side of Fig. 5.8. The incident radiation $b(\phi)$ is sampled at

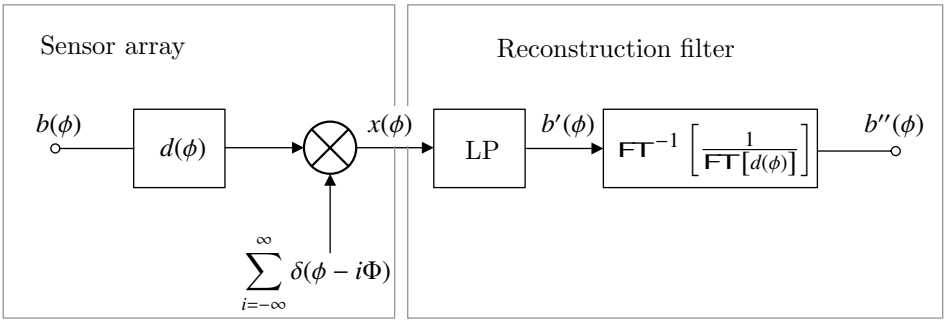


Figure 5.8: Signal flow of the angle of arrival detector

spaces of Φ with respect to the sensitivity profile $d(\phi)$. The resulting signal is

$$x(\phi) = (b(\phi) * d(\phi)) \cdot \sum_{i=-\infty}^{\infty} \delta(\phi - i\Phi), \tag{5.4}$$

where the operator $*$ denotes convolution, and $\delta(\phi)$ is Dirac's delta function. The radiation signal $b(\phi)$ can be reconstructed from $x(\phi)$ using a low pass filter and a sensitivity pattern compensation filter. For perfect reconstruction, the following conditions must be met:

- For the maximum frequency of $b(\phi)$, $f_{b,max} < \frac{1}{2\Phi}$ must hold (Nyquist-Shannon sampling theorem).
- The Fourier transform of the sensitivity profile, $\mathrm{FT}\left[d(\phi)\right]$, must not contain any roots within $\left[0, \frac{1}{2\Phi}\right)$.

With an ideal low pass filter defined by the transfer function[2]

$$H_{LP}(f) = \mathrm{rect}\left(f \cdot \Phi\right),$$

the interpolated signal $b'(\phi)$ is given by

$$b'(\phi) = \frac{1}{\Phi}\,\mathrm{si}\left(\frac{\pi\phi}{\Phi}\right) * x(\phi).$$

Finally, using the low pass filter and a sensitivity profile compensation filter, the reconstructed radiation profile, $b''(\phi)$, is

$$b''(\phi) = \mathrm{FT}^{-1}\left[X(f) \cdot H_{LP}(f) \cdot \frac{1}{\mathrm{FT}\left[d(\phi)\right]}\right],$$

[2]The rect function is defined in Eq. 5 on p. XII.

where $X(f)$ is the Fourier transform of $x(\phi)$. If the above mentioned conditions are met, then $b''(\phi) = b(\phi)$ holds, i.e. the radiation profile is perfectly reconstructed. Please note that the prime signs $'$ in $b'(\phi)$ and $b''(\phi)$ do *not* indicate the first and second derivative.

The dominant angles of radiation can be determined by finding the absolute maximum of $b''(\phi)$. Hence, for adequate frequency signals, i.e., signals that meet the above conditions, a very high AOA resolution can be achieved.

In the real world it is unlikely that at all times the sampling theorem is met. Instead, the incident radiation profile will contain frequency components that are not perceived by the detector. As a consequence, the radiation signal cannot unambiguously be reconstructed due to aliasing. This situation can be prevented by adding more sensors and thus increasing the sampling rate. However, for some signal types such as extremely narrow beams, this would result in a very high number of sensors. Alternatively, an *approximation* of the true angle of arrival may in many applications turn out to be sufficient. Therefore the optimum detector design is based on a trade off between the AOA estimation error and the number of sensors, i.e. detector complexity.

5.6 Practical approach using numerical analysis

We have now described the general theoretical idea of the AOA detector. Since the sensor data are most easily processed digitally we continue investigations in the discrete domain. The (continuous) profile functions $b(\phi)$, $b'(\phi)$, $b''(\phi)$ and the sensor sensitivity profile $d(\phi)$ are then represented by sequences of N samples in the angular domain, i.e. by $\mathbf{b}^{\mathrm{T}} = \{b(n)\}$, $\mathbf{b'}^{\mathrm{T}} = \{b'(n)\}$, $\mathbf{b''}^{\mathrm{T}} = \{b''(n)\}$ and $\mathbf{d}^{\mathrm{T}} = \{d(n)\}$, resp. With Δr as sampling resolution we define

$$\begin{aligned} N &= \left\lfloor \frac{360^\circ}{\Delta r} \right\rfloor, \\ n &= \left\lfloor \frac{180^\circ + \phi}{\Delta r} \right\rfloor \bmod N, \end{aligned} \tag{5.5}$$

with $0 \le n < N$. The changeover to discrete angles is legitimate if the sampling theorem,

$$\begin{aligned} f_{b,max} &< \frac{1}{2\,\Delta r} \quad \text{and} \\ f_{d,max} &< \frac{1}{2\,\Delta r}, \end{aligned}$$

holds, where $f_{d,max}$ and $f_{b,max}$ are the upper angular frequency bounds of the sensor sensitivity and incident radiation profiles. For all following considerations, we assume this to be the case. Also, in the examples to follow, a resolution of $\Delta r = 0.5^\circ$ is chosen. Therefore, all signals are described by sequences of length $N = 720$. Since N is sufficiently large, a signal is plotted continuously, not at single sampling points. The frequency spectrum is mirrored at intervals

of $\frac{1}{\Delta r}$, or N samples. With (5.5), the frequency resolution (measured in 1/1°) is hence given by

$$\Delta f = \frac{1}{N \cdot \Delta r} = \left(\left\lfloor \frac{360^\circ}{\Delta r} \right\rfloor \cdot \Delta r \right)^{-1}. \tag{5.6}$$

5.6.1 Further signal representations in the discrete signal space

The sensitivity profile of the l-th sensor $d_l(n)$ can now be described by a vector $\mathbf{d}_l = \{d_l(n)\}^{\mathrm{T}}$. Note that the resulting sequence takes into account the angular displacement ρ_l of the respective sensor. Using Eq. 5.1, and according to Eq. 5.3, its elements are given by

$$d_l(n) = \begin{cases} \cos^\kappa \left((n \cdot \Delta r - \rho_l - 180^\circ) \cdot \frac{90^\circ}{\tau} \cdot \frac{\pi}{180^\circ} \right), & \text{if } -\tau \leq n \cdot \Delta r - \rho_l - 180^\circ < \tau \\ 0, & \text{otherwise.} \end{cases} \tag{5.7}$$

In the next step, we merge all L vectors $\mathbf{d}_l$ to the $(L \times N)$ sensitivity pattern matrix $\underline{\mathbf{D}}$, such that every entry corresponds to the sensitivity of a certain sensor at one particular angle of incidence. For instance, $d_0(N/2)$ represents the sensitivity of the leftmost sensor (indexed by zero) in the center of the angular range, $N/2$, i.e. 0°.

$$\underline{\mathbf{D}} = \begin{bmatrix} \mathbf{d}_0^{\mathrm{T}} \\ \mathbf{d}_1^{\mathrm{T}} \\ \vdots \\ \mathbf{d}_{L-1}^{\mathrm{T}} \end{bmatrix} = \underbrace{\begin{bmatrix} d_0(0) & d_0(1) & \dots & d_0(N-1) \\ d_1(0) & d_1(1) & \dots & d_1(N-1) \\ \vdots & \vdots & \ddots & \vdots \\ d_{L-1}(0) & d_{L-1}(1) & \dots & d_{L-1}(N-1) \end{bmatrix}}_{N \text{ angles}} \Bigg\} L \text{ sensors} \tag{5.8}$$

The L sensor readings x_l can be expressed by the product of the sensitivity patterns of the sensors, $d_l(n)$, and the radiation profile, $b(n)$, resulting in the vector $\mathbf{x}$:

$$\mathbf{x} = \{x_l\}^{\mathrm{T}} = \underline{\mathbf{D}}\,\mathbf{b} \tag{5.9}$$

According to the basic approach introduced in Sec. 5.5, the (discrete) original signal can be reconstructed from these sensor readings using a low pass interpolation filter and a sensitivity pattern compensation filter. Using Eq. 5.6, the cut-off frequency f_c of the interpolation filter is defined by

$$f_c = \frac{360^\circ}{2\Phi} \Delta f.$$

With $N_c = \lfloor f_c/\Delta f \rfloor$, the transfer function $H_{LP}(n)$ is given by

$$H_{LP}(n) = \begin{cases} 1, & \text{if } 0 \le n < N_c \\ 0, & \text{if } N_c \le n < N - N_c \\ 1, & \text{if } N - N_c \le n < N. \end{cases} \tag{5.10}$$

At this moment, $\mathbf{x}$ is a vector of length L, but the range of angles has width N. Before we can interpolate the sensor readings, $\mathbf{x}$ must be upsampled (i.e., stuffed with zeros) to fit length N. Subsequently, the upsampled sensor readings $\tilde{x}(n)$ are convolved with the impulse response $h_{LP}(n) = \mathrm{DFT}^{-1}[H_{LP}(n)]$, where $\otimes$ indicates cyclic convolution:

$$b'(n) = \tilde{x}(n) \otimes h_{LP}(n) = \sum_{m=0}^{N-1} \tilde{x}(m) \cdot h_{LP}((n-m) \bmod N) \tag{5.11}$$

The resulting signal, $b'(n)$, represents an approximation of the radiation signal $b(n)$, as shown in Fig. 5.9.

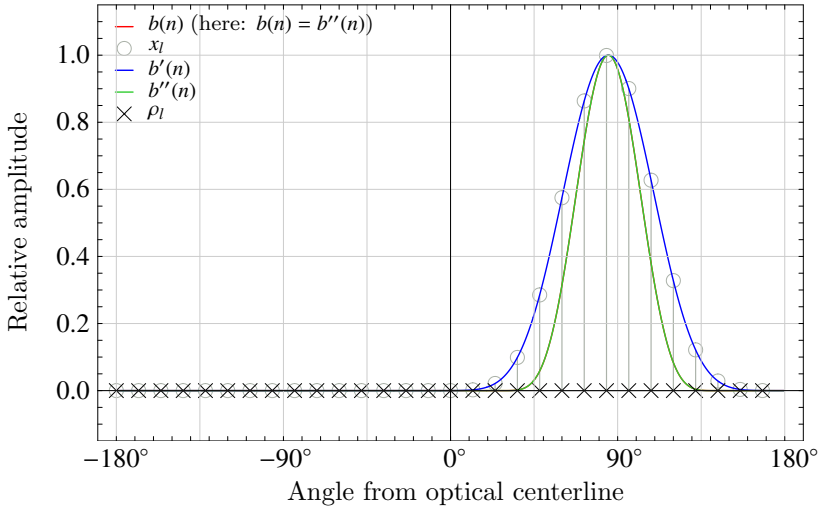


Figure 5.9: Signal recovery by sample interpolation and sensitivity profile compensation in the case of no alias. $L = 30$, $\kappa = 1$, $\tau = 45°$

However, $b'(n)$ does not yield perfect signal recovery. As shown in the continuous signal space, the sensitivity profile compensation filter must additionally be applied. Following the operational order as shown in Fig. 5.8, we would first upsample the sensor readings x_l (as shown above), perform interpolation, and then compensate the sensitivity profile. The latter operation would be a convolution of sequences of length N. However, if $f_{d,max} < \frac{1}{2\Phi}$ holds, the sensitivity profile $d(n)$ can be expressed by L (instead of N) samples without loosing information. We denote the elements of this profile $d_s(l)$, using Eq. 5.7:

$$d_s(l) \;=\; d_0\left(\frac{\rho_l + 180°}{\Delta r}\right), \quad 0 \le l < L$$

For reasons of lower complexity, we can now exchange the low pass interpolation operation (which requires upsampling) and the sensitivity profile compensation. I.e., first the sensor readings x_l are filtered with the sensitivity profile compensation filter (which is now a convolution of sequences of length L) and subsequently, the result is interpolated using the low pass filter. For fast convolution, the sensor readings x_l and the sensitivity profile $d_s(l)$ are transformed into the spectral domain and element-wise divided by each other[3]:

$$e(l) \;=\; \mathrm{DFT}^{-1}\left[\frac{\mathrm{DFT}\,[x_l]}{\mathrm{DFT}\,[d_s(l)]}\right]$$

The inverse transformed result $e(l)$ is upsampled – such that it becomes a sequence of N samples, denoted $\tilde{e}(n)$ – and finally convolved with the low pass impulse response $h_{LP}(n)$:

$$b''(n) = \tilde{e}(n) \otimes h_{LP}(n)$$

The result of these operations is illustrated in Fig. 5.9. In this case, all conditions for perfect reconstruction are met, such that $b''(n) = b(n)$.

Apparently, the impact of the sensitivity profile compensation filter is marginal: it only slightly bulges the shape of the signal. The angular position of the peak – which is the information of interest – does not shift. The reason for this behavior can be found in Eq. 5.1 and Eq. 5.4: $b(\phi)$ is convolved with an *even* signal (i.e., $d(\phi) = d(-\phi)$) with its maximum at $\phi = 0$. Therefore, no angular displacement of the peak of $b(\phi)$ takes place[4]. As a consequence, the compensation of the sensitivity profile can be omitted. In the following, we consider only the interpolated samples approach, yielding $b'(n)$. The sensitivity profile compensation filter can be omitted.

5.6.2 AOA error in the case of alias

In the example above it was assumed that no aliasing takes place. In the real world, this situation is highly unlikely, especially when receiving narrow *beams* of radiation. Fig. 5.10 shows the result of the interpolation filter in the case of alias for differently shaped radiation signals $b(n)$. The absolute AOA error of the reconstructed signal $b'(n)$ is denoted $\epsilon_{b'}(\alpha)$ and is given by

$$\epsilon_{b'}(\alpha) = |\,\alpha - \alpha_{b'}\,|, \tag{5.12}$$

where α and $\alpha_{b'}$ indicate the *true* and the *measured* angle of arrival, respectively:

$$\alpha = \mathrm{argmax}\,(b(\phi))\,, \quad \alpha_{b'} = \mathrm{argmax}\left(b'(\phi)\right)$$

In the example shown in Fig. 5.10(a), the true angle of arrival of a narrow beam is $\alpha = -35°$, whereas the recovered signal has its peak at $\alpha_{b'} = -39°$. Contrary to this, with the broader

[3]Note that for this operation, $\mathrm{DFT}\,[d_s(l)]$ must not contain any roots.

[4]Also, $b(\phi)$ is assumed to be symmetric around its peak angle.

(a) Narrow radiation signal, $\alpha = -35°$, $\alpha_{b'} = -39°$ (b) Wide radiation signal, $\alpha = -35°$, $\alpha_{b'} = -35°$

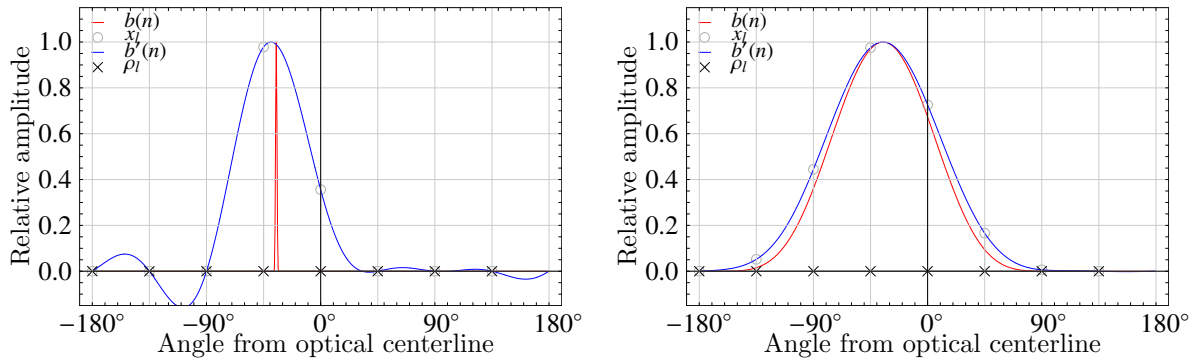


Figure 5.10: Signal recovery by sample interpolation and sensitivity profile compensation, in the case of alias. $L = 8$, $\kappa = 1$, $\tau = 45°$

radiation profile shown in Fig. 5.10(b) the identical detector yields the exact angle of arrival. This is due to the lack of high frequency components and therefore less alias in the latter case. This example shows the impact of the shape of the radiation profile.

Besides the radiation profile, the angle of incidence itself determines the magnitude of the measurement error. Fig. 5.11 shows the periodicity of the error against the true angle of arrival. The radiation profile applied in this example was chosen according to the narrow beam in Fig. 5.10(a). The beam was swept over the entire angular range.

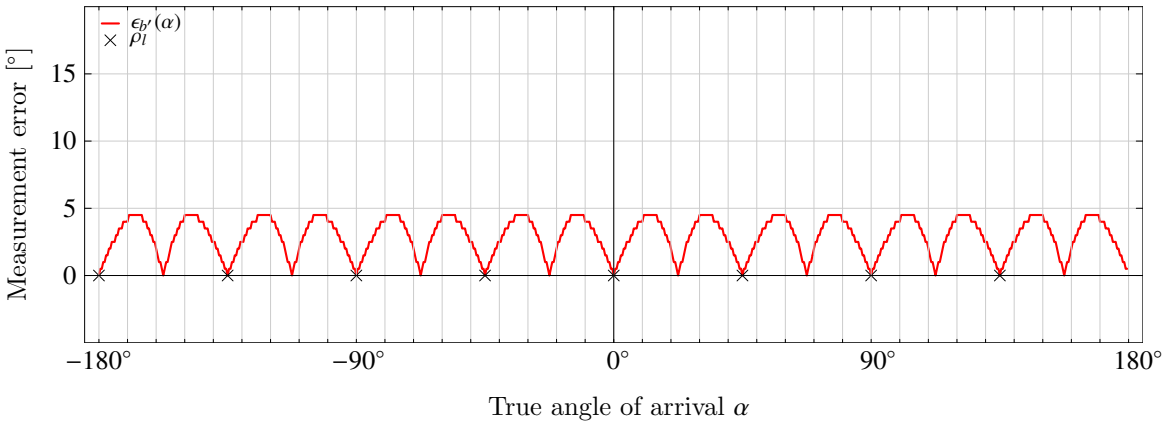


Figure 5.11: Absolute measurement error vs. angle of incidence. $L = 8$, $\Phi = 45°$, $\kappa = 1$, $\tau = 45°$

5.6.3 Computational effort

The interpolation sequence $\{b'(n)\}$ is the result of a cyclic convolution on real data $\{\tilde{x}(n)\}$ and $\{h_{LP}(n)\}$, as defined in Eq. 5.11. The arithmetic complexity of this operation is given by N^2 multiplications and $N(N-1)$ additions. In the following, the number of multiplications and additions are denoted by the letters M and A, respectively. Since $\{\tilde{x}(n)\}$ is zero for all but L elements, the effort can be drastically reduced. Furthermore, if $\{h_{LP}(n)\}$ is an even sequence[5] then only $N/2+1$ multiplications must be computed. This leads to the following number of operations:

$$M_N^{\text{conv direct}} = (N/2+1) \cdot L \tag{5.13}$$

$$A_N^{\text{conv direct}} = N \cdot (L-1) \tag{5.14}$$

In practical implementations usually a fast convolution based on the fast Fourier transform (FFT) is favored (Fig. 5.12). When using a split radix FFT and N is a power of 2, the

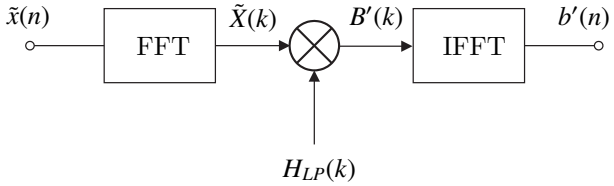


Figure 5.12: Fast convolution using the FFT

complexity is

$$\begin{aligned} \tilde{M}_N^{\text{conv FFT}} &= \overbrace{2\,M_N^{\text{RFFT}}}^{\text{FFT \& IFFT}} + \overbrace{2+3\,(N/2-1)}^{\text{spectrum mult.}} \\ &= N/2\,(2\,\log_2 N - 3) + 3, \end{aligned}$$

$$\begin{aligned} \tilde{A}_N^{\text{conv FFT}} &= \overbrace{2\,A_N^{\text{RFFT}}}^{\text{FFT \& IFFT}} + \overbrace{3\,(N/2-1)}^{\text{spectrum mult.}} \\ &= N/2\,(6\,\log_2 N - 7) + 5, \end{aligned}$$

where RFFT denotes an FFT on real data or an IFFT of complex data with hermitian symmetry ($\tilde{X}(k) = \tilde{X}^*(N-k)$) [25, 82]. These two operations have the same arithmetic complexity. Due to the symmetry of both $\{\tilde{X}(k)\}$ and $\{H_{LP}(k)\}$, $\{B'(k)\}$ can be obtained with two real multiplications for $k = 0$ and $k = N/2$, and $N/2-1$ complex multiplications for $k = 1 \ldots N/2-1$. In the context of the FFT, the latter can be split into 3 real multiplications and 3 real additions each [82][6].

[5] A sequence $\{a(n)\}$ is even if $a(i) = a(N-i)$ holds for all $i \in \{0 \ldots N-1\}$.

[6] Basically, a multiplication of two complex numbers can be decomposed to 4 real multiplications and 2 real additions, or, alternatively, to 3 real multiplications and 5 real additions. In the context of the FFT, 2 of these 5 additions can be precalculated, such that only 3 real multiplications and 3 real additions remain.

A few simplifications can be achieved that reduce the complexity. In Eq. 5.10 $\{H_{LP}(k)\}$ has been defined to be entirely real. It even consists only of ones and zeros. As a result, no complex multiplications need be performed at all[7].

Furthermore, a large number of trivial zero-multiplications occur since $\tilde{x}(n)$ is non-zero only at L positions. As the position and the amount of sensors may be unknown beforehand[8], it cannot clearly be decided how many trivial multiplications emerge: the FFT is performed in $\log_2 N$ stages. In each stage, all elements of the sequence to be transformed are in-place substituted by a new value. This value is affected by the result of a complex multiplication. In each stage, only $N/2$ such multiplications (the multipliers being either the even or odd indexed elements) are performed. Hence, the position of the L non-zero elements in the sequence $\{\tilde{x}(n)\}$ is relevant. In the worst case, if all sensors refer to even (or odd) indexed sequence positions and in the first stage, all multiplications are performed on odd (or even) indexed elements, then no trivial multiplications occur. Therefore, in this complexity estimation, the rate of trivial operations is ignored. The computational complexity of the low pass interpolation approach using the FFT is thus

$$\begin{aligned} M_N^{\text{conv FFT}} &= 2\,M_N^{\text{RFFT}} \\ &= N(\log_2 N - 3) + 4 \end{aligned} \tag{5.15}$$

$$\begin{aligned} A_N^{\text{conv FFT}} &= 2\,A_N^{\text{RFFT}} \\ &= N/2\,(6\,\log_2 N - 10) + 8 \end{aligned} \tag{5.16}$$

Note that the split radix FFT algorithm applied here has its strengths in its rather low arithmetic complexity. The disadvantage of this scheme is that it needs both forward transform on real data and inverse transform on complex data with conjugate complex symmetry. However, it was chosen here because of its broad coverage in the literature. Other arithmetically less efficient but structurally advantageous schemes like the discrete Hartley transform (DHT) might be an alternative for practical implementation [25].

Interpolation by cubic splines – although being mentioned in Sec. 5.6.5 – is not a subject of this section. The computational effort is significantly larger than that of low pass convolution.

To compare the two convolution methods, it is reasonable to consider only the number of multiplications since this operation is considerably more complex than an addition. Comparing Eq. 5.13 and 5.16, direct convolution yields a better algorithmic efficiency if

$$\begin{aligned} M_N^{\text{conv direct}} &< M_N^{\text{conv FFT}} \\ \left(\frac{N}{2} + 1\right) L &< N\left(\log_2 N - 3\right) + 4 \end{aligned}$$

[7] A complex number z multiplied by 1 or 0 results in $z \cdot 1 = z$ and $z \cdot 0 = 0$, resp.

[8] It is assumed that different device configurations – and hence, number of sensors – can be built using the very same fast convolution algorithm. If the algorithm is adapted to the sensor arrangement of each detector array individually, a further reduction of complexity can be obtained.

holds. This inequation can be written as

$$L < \frac{2N\left(\log_2 N - 3\right) + 8}{N + 2},$$

and because $2N\left(\log_2 N - 3\right) \gg 8$ and $N \gg 2$ we get the approximation

$$L < 2\log_2 N - 6.$$

E.g., for $N = 256$, direct convolution requires less computational effort if $L < 10$. The fact that direct convolution can be the more efficient algorithm seems contradicting since usually fast convolution exploiting the FFT is significantly faster. The superiority of the direct convolution algorithm is due to the large number of simplifications we could accomplish. Some of these cannot be applied to the fast convolution method offhand, and thus the number of operations cannot further be decreased.

5.6.4 Field of view

Unless mounted on a mobile device, the AOA detector is mounted on walls. In this case, only a certain field of view must be covered since only within a certain angular range, radiation for angle of arrival detection is expected. For instance, when mounted in a rectangular corner, the field of view is limited to 90°. The principle design of this modified IR AOA detector array is shown in Fig. 5.13. It consists of L narrowly placed sensors that are angularly displaced against each other[9]. The amount of rotation between two sensors is denoted Φ. The field of view (FOV), i.e. the angular range of the detector array which is exposed to radiation, is constrained by delimiters at angles $\pm\gamma_{FOV}$.

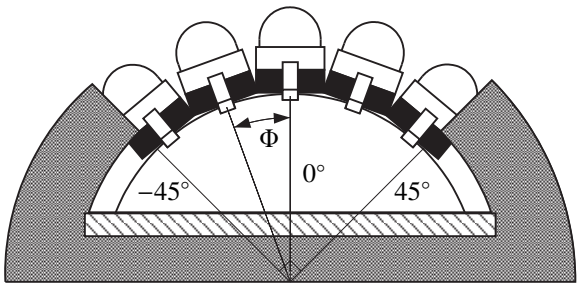


Figure 5.13: IR detector array with a 90° field of view ($\gamma_{FOV} = 45°$), $L = 5$, $\Phi = 20°$

The sensor rotations ρ_l are now given by

$$\rho_l = -\frac{\Phi}{2}(L - 2l - 1).$$

[9]More precisely, we assume that all sensors are placed at the same location and are only rotated against each other. However, the approximated arrangement shown in Fig. 5.13 is permitted since the distance between two sensors is negligible compared to the distance of the radiation source.

For each sensor $l \in \{0, \ldots, L-1\}$, a shifted sensitivity profile $d_l(n)$ can be determined according to Eq. 5.7. The interpolation method for signal recovery can be applied as described in the previous section. Since the angular range – defined by the field of view – is not seamlessly periodic anymore, the AOA measurement error $\epsilon_{b'}(\alpha)$ rises towards the edges of the field of view. This behavior is illustrated in Fig. 5.14 for detector arrangements consisting of four (a), five (b), seven (c) and nine (d) sensors, respectively. Each diagram can roughly be divided

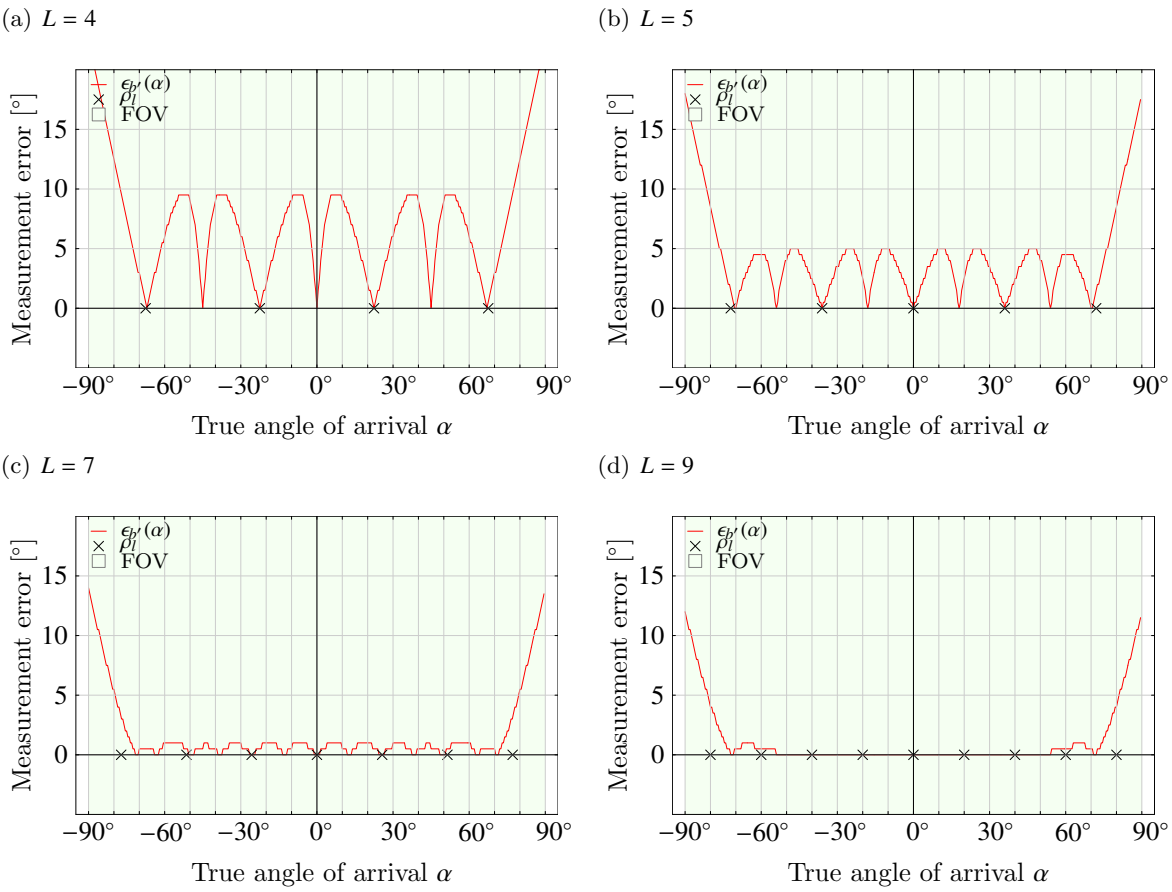


Figure 5.14: Absolute measurement error vs. angle of incidence. $\kappa = 10$, $\tau = 90°$, $\rho_l = -\frac{\gamma_{FOV}}{L}(L - 2l - 1)$, $\gamma_{FOV} = 45°$

into three parts, a *central part* ($\rho_0 \leq \alpha \leq \rho_{L-1}$), and two *border parts* ($-\gamma_{FOV} < \alpha < \rho_0$ and $\rho_{L-1} < \alpha < \gamma_{FOV}$, respectively). In fact the error in the latter is immensely higher than it is in the central part. In all parts, the deviation from the true angle decreases with rising L. The mean AOA measurement error $\langle \epsilon_{b'}(L) \rangle$ for an increasing number of sensors L is shown in Fig. 5.15. Apparently, the average error cannot be decreased below 1°. However, when regarding Fig.5.14(d) it becomes obvious that this result is degraded by the border parts. In a large central part (from −53° to 53°) the error is zero. Therefore, we will now investigate the impact of altering the angular sensor displacement.

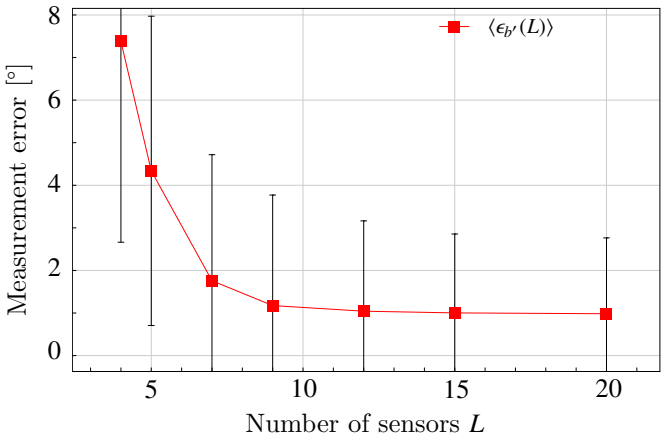


Figure 5.15: Mean $\langle \epsilon_{b'}(L) \rangle$ and standard deviation $\sigma_{b'}(L)$ of measurement error vs. number of sensors L

5.6.5 Unbalanced sensor orientation

In some applications, a constant measurement accuracy is desired. In other cases, it may be advantageous to achieve a significantly higher accuracy towards 0° than at higher angles. This is for example reasonable when installing the detector at the head of a long narrow hallway. Therefore, an application specific distribution of sensors may lead to the most beneficial result. Fig. 5.16 shows the error for detector arrays of seven differently arranged sensors. In the first case 5.16(a) the sensors are displaced by equal amounts of $\Phi = 2\,\gamma_{FOV}/(L-1)$, i.e., less dense[10] than in Fig. 5.16(b). As a result, the error in the former central part increases. Also, the immense error at the border parts has vanished. In the second case 5.16(b) the sensor density is higher in the central part than towards the borders. Since the sensor readings do no longer represent equally distant samples, a cubic spline algorithm instead of a low pass filter was used for interpolation. The measurement resolution in the sector spanned by ±33° has increased in return for the decreased accuracy in the remaining field of view. Note that some of these observations strongly depend on the sensitivity profile of the sensors. Hence, with a different type of sensors, totally different results may be obtained.

5.7 Sensitivity aberrations

Real world detector arrays are subject to random and systematic aberrations concerning the sensitivity characteristics. Fig. 5.17 shows the deviations of the single sensor sensitivity profiles of a thermopile array (from [44])[11]. Also, the electro-optical characteristics of single

[10] in the angular sense

[11] Thermopiles are thermal detectors briefly discussed in Sec. 6.4. The angular sensitivity properties of these devices are similar to those of photodetector arrays.

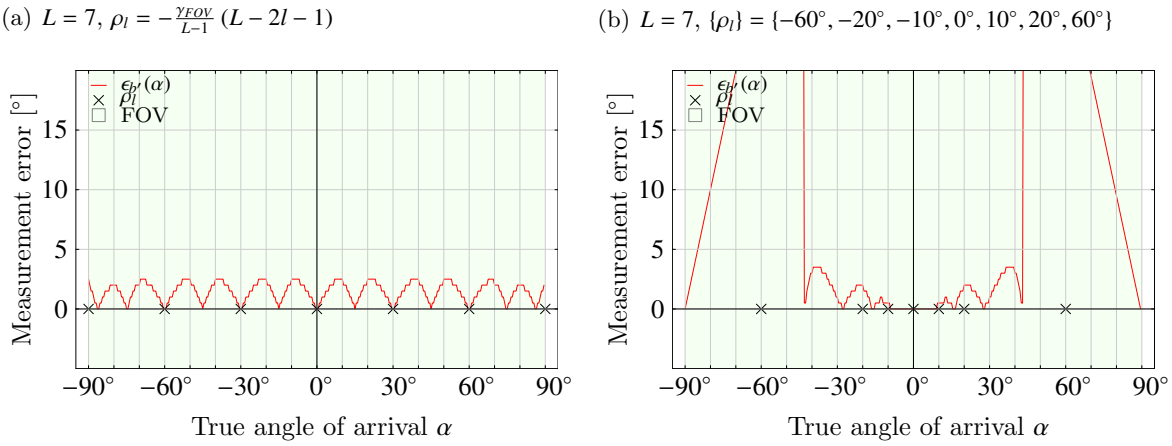


Figure 5.16: Absolute measurement error vs. angle of incidence for an alternative placement of sensors. $\kappa = 10$, $\tau = 90°$

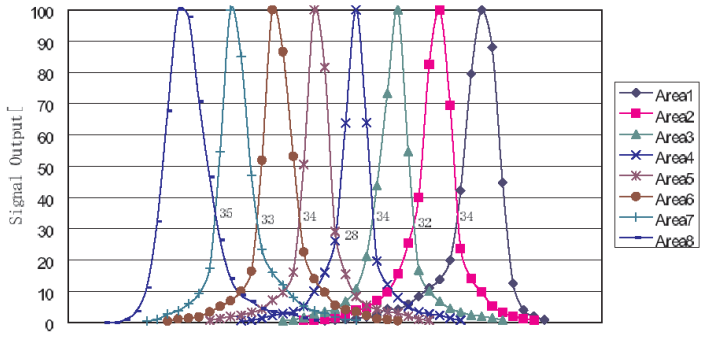


Figure 5.17: Sensitivity profile of an 8 element thermopile array

sensor elements can deviate significantly from the values given in their datasheet. Iske et al. carried out experiments to analyze this issue on photodetectors [48]. The angular characteristics as stated in the datasheet did not comply with the real measurements. Different angles of half sensitivity[12] were observed. Moreover, the measured sensitivity profile turned out to be non-symmetric in one case.

According to these examples the variance of the sensor characteristics must be taken into account. However, in the previous sections, it has been assumed that all sensors' properties are identical. As a matter of fact, equality is a precondition for the signal recovery method described in the previous sections. Particularly, interpolation using a low pass filter is critical if the radiation profile is not sampled at constant intervals. Hence the measurement precision of the detector decreases significantly without further modifications.

To account for the variance in device sensitivity we extend Eq. 5.7 to

$$d_l(n) = \begin{cases} (1 + a_l \cdot r_l(n)) \cdot \cos^{\kappa_l}\left((n\,\Delta r - \rho_l - 180°) \cdot \frac{90°}{\tau_l} \cdot \frac{\pi}{180°}\right), & \text{if } (N_- - n) \bmod N \le 0 \\ & \quad \wedge\ 0 < (N_+ - n) \bmod N \\ 0, & \text{otherwise.} \end{cases} \tag{5.17}$$

where

$$N_\pm = \left(\frac{\rho_l \pm \tau_l}{\Delta r} + \frac{N}{2}\right) \bmod N$$

and

$$\kappa_l = \kappa_0 + \Delta\kappa_l, \quad \tau_l = \tau_0 + \Delta\tau_l \quad \text{and} \quad \rho_l = \rho_{l,0} + \Delta\rho_l.$$

$\Delta\kappa_l$, $\Delta\tau_l$ and $\Delta\rho_l$ are the absolute deviations from the respective nominal values κ_0, τ_0 and $\rho_{l,0}$ at sensor l and are considered Gaussian, with mean 0 and standard deviation σ_κ, σ_τ, σ_ρ. They are caused by mounting and manufacturing tolerances or due to construction constraints. The first two affect the shape of the sensor's sensitivity profile, whereas the latter determines the angular displacement inaccuracy of the sensor. The sinusoidal term in Eq. 5.17, $\cos^{\kappa_l}(\ldots)$, is modulated by an error function $r_l(n)$. This may e.g. be made up of low frequency components to add an asymmetry to the sensitivity profile. The modulation intensity is indicated by weighting factor a_l which is determined by a Gaussian process with mean zero and standard deviation σ_a. Heuristically the following error function $r_l(n)$ was chosen:

$$r_l(n) = \begin{cases} \cos\left(4 \cdot \frac{\pi}{\tau_l} \cdot (n\,\Delta r + \nu_l)\right), & \text{if } (N_- - n) \bmod N \le 0 \ \wedge\ 0 < (N_+ - n) \bmod N \\ 0, & \text{otherwise.} \end{cases} \tag{5.18}$$

ν_l is generated by a random process and uniformly distributed in the interval $[0, 2\tau_l)$.

An example set of sensitivity profiles using this error function definition is shown in Fig. 5.18.

[12] The angle of half sensitivity is defined as the angle of incidence for which the sensitivity of the receiver has decreased to 50% of the peak sensitivity.

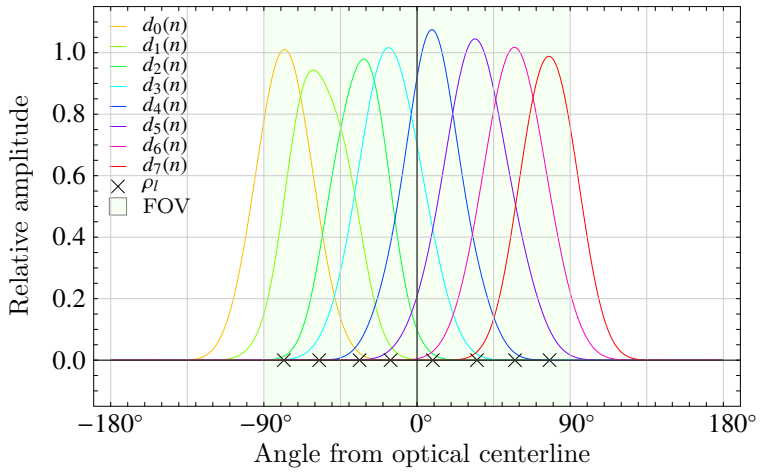


Figure 5.18: Deviated sensitivity profiles ($L = 8$, $\sigma_\kappa = 0$, $\sigma_\tau = 5°$, $\sigma_\rho = 2°$, $\sigma_a = 0.05$)

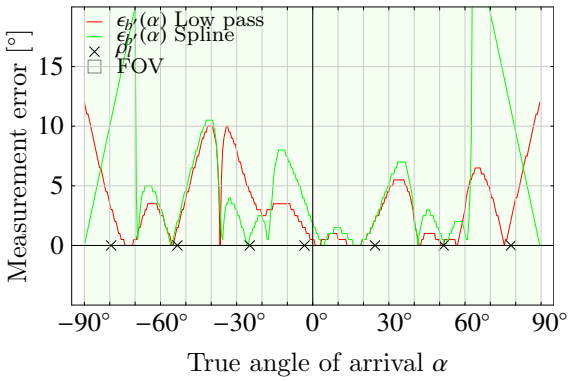


Figure 5.19: Measurement error using low pass and cubic spline interpolation for an irregular sensor array. $\kappa_0 = 10$, $\tau_0 = 90°$, $\rho_{0,l} = -\frac{\gamma_{FOV}}{L}(L - 2l - 1)$

Signal reconstruction by interpolation as presented in the previous sections is strongly affected by these irregularities. However, some of these can be compensated. For this purpose, each sensor's sensitivity profile can be gauged in a calibration measurement. Unfortunately, interpolation by low pass requires the sensors to be equally distributed in the angular range, i.e. $\Delta\rho_l = 0$. Therefore, the outcome of the method strongly depends on the precise mounting and/or optics of the sensor array. Furthermore, all sensors' sensitivity profiles must be identical ($\Delta\kappa_l = const$, $\Delta\tau_l = const$). This latter requirement is also true if spline interpolation is performed. In this case, the unbalanced orientation of the sensors ($\rho_l \neq 0$) can be compensated by calibration measurements. But, since the sensor readings are merely treated as samples, the irregular and differing sensitivity profiles are not taken into account and hence tamper the result. Fig. 5.19 shows the reconstruction error of an array of seven sensors which are misaligned by $\sigma_\rho = 2°$. Its sensitivity profiles are distorted according to Eq. 5.17 and 5.18, with $\sigma_\kappa = 0$, $\sigma_\tau = 15°$ and $\sigma_a = 0.1$. Comparing these results with the outcome of the ideal

detector in Fig. 5.14(c), the impact of detector irregularities becomes obvious.

5.8 Signal recovery by virtual filter interpolation (VFI)

The results of the interpolation techniques discussed above, low pass and spline, yields the question for a superior approach that takes all sensor sensitivity irregularities into account. In this section, a novel technique for angle of arrival detection using an IR detector array is presented. It is based on the idea that a digital filter is emulated by combining the sensitivity profiles of the individual sensors in a certain way. It requires a one-time calibration measurement of the individual sensor sensitivities, as has been mentioned in the previous section. But contrary to the interpolation technique, not only the orientation offset of each sensor, but the complete sensor sensitivity profile is taken into account. The method was named *virtual filter interpolation* (VFI).

5.8.1 The filter emulation approach

The signal flow of the VFI detector array can be modeled as shown in Fig. 5.20. The L sensor

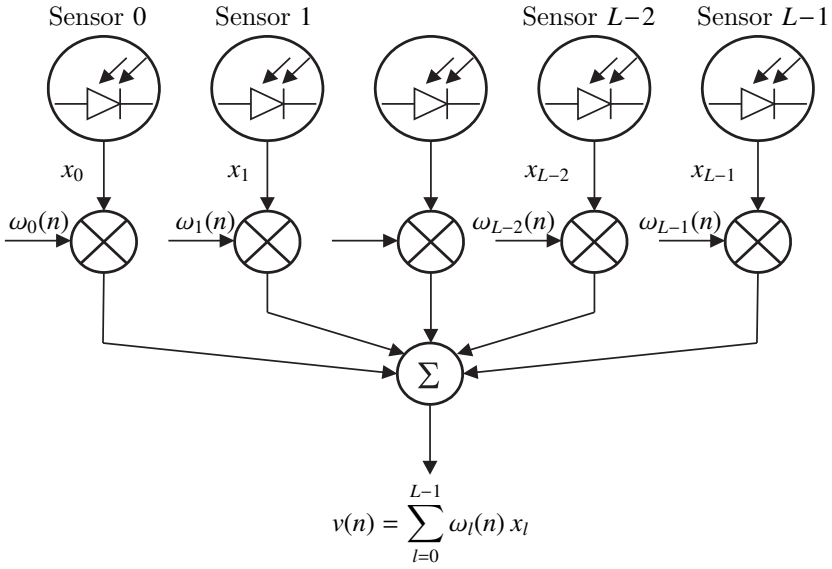


Figure 5.20: Signal flow graph of the VFI detector

readings x_l are weighted by coefficients $\omega_l(n)$ before being summed up to $v(n)$. For every angle $\phi = (n - N/2)\,\Delta r$ the sensor readings are weighted differently. The sensor weights $\omega_l(n)$ can be any real number and therefore the impact of certain selected sensors can be emphasized or expunged totally.

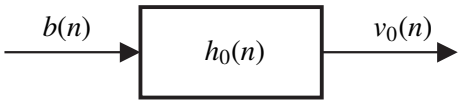


Figure 5.21: Virtual angle of arrival filter

Consider we had N sets of weighting coefficients $\omega_n(l)$ that are applied sequentially, written as weighting matrix $\underline{\mathbf{\Omega}}$:

$$\underline{\mathbf{\Omega}} = \begin{bmatrix} \omega_0(0) & \omega_0(1) & \dots & \omega_0(L-1) \\ \omega_1(0) & \omega_1(1) & \dots & \omega_1(L-1) \\ \vdots & \vdots & \ddots & \vdots \\ \omega_{N-1}(0) & \omega_{N-1}(1) & \dots & \omega_{N-1}(L-1) \end{bmatrix} = \left.\begin{bmatrix} \boldsymbol{\omega}_0^{\mathrm{T}} \\ \boldsymbol{\omega}_1^{\mathrm{T}} \\ \vdots \\ \boldsymbol{\omega}_{N-1}^{\mathrm{T}} \end{bmatrix}\right\} N \text{ sets of } L \text{ sensor weights}$$

Then, with $\underline{\mathbf{\Omega}}$ and Eq. 5.8 and 5.9, and introducing $\underline{\mathbf{H}} = \underline{\mathbf{\Omega}}\underline{\mathbf{D}}$, the resulting vector $\mathbf{v} = \{v(n)\}^{\mathrm{T}}$ becomes

$$\mathbf{v} = \underline{\mathbf{\Omega}}\,\mathbf{x} = \underline{\mathbf{\Omega}}\,\underline{\mathbf{D}}\,\mathbf{b} = \underline{\mathbf{H}}\,\mathbf{b}. \tag{5.19}$$

Unfortunately, $\underline{\mathbf{H}}$ is a singular matrix and hence, this equation system cannot uniquely be solved for $\mathbf{b}$: $\underline{\mathbf{D}}$ has rank L, and the rank of the product of two matrices satisfies the relation $\mathrm{rank}(\underline{\mathbf{H}}) \leq \min(\mathrm{rank}(\underline{\mathbf{\Omega}}), \mathrm{rank}(\underline{\mathbf{D}}))$. Independent of the rank of $\underline{\mathbf{\Omega}}$, $\mathrm{rank}(\underline{\mathbf{H}}) \leq L$ holds. Since $\underline{\mathbf{H}}$ is an $N \times N$ matrix, and $L < N$, $\underline{\mathbf{H}}$ is singular and therefore $\underline{\mathbf{H}}^{-1}$ does not exist.

This result is as anticipated, since $\mathbf{x}$ lacks relevant information to reconstruct $\mathbf{b}$ precisely. So how can a reasonable solution nevertheless be found? In the VFI approach, a number of constraints are defined that must be met as well as possible by the system. Then, with signal processing means, an approximated radiation signal can be reconstructed. We introduce $\underline{\mathbf{H}}_0$ and $\mathbf{v}_0 \approx \mathbf{v}$ with

$$\mathbf{v}_0 = \underline{\mathbf{H}}_0\,\mathbf{b}. \tag{5.20}$$

Provided that $\underline{\mathbf{H}}_0$ is a circulant matrix of the form

$$\underline{\mathbf{H}}_0 = \begin{bmatrix} h_0(0) & h_0(N-1) & \dots & h_0(1) \\ h_0(1) & h_0(0) & \dots & h_0(2) \\ \vdots & \vdots & \ddots & \vdots \\ h_0(N-1) & h_0(N-2) & \dots & h_0(0) \end{bmatrix}, \tag{5.21}$$

the matrix vector product $\underline{\mathbf{H}}_0\,\mathbf{b}$ in Eq. 5.20 can be interpreted as a cyclic convolution, and $h_0(n)$ may be called *impulse response* of the discrete LSI (linear, shift invariant) system shown in Fig. 5.21 (e.g. [68]). The output signal of this filter, $v_0(n)$, is given by

$$v_0(n) = b(n) \otimes h_0(n) = \sum_{m=0}^{N-1} b(m) \cdot h_0\left((n-m) \bmod N\right). \tag{5.22}$$

According to Eq. 5.19, $\underline{\mathbf{H}}$ is composed of the set of weighting coefficients, $\underline{\mathbf{\Omega}}$, and the set of sensor sensitivity profiles, $\underline{\mathbf{D}}$. Our goal is to choose $\underline{\mathbf{\Omega}}$ in a way that $\underline{\mathbf{H}}$ approximates $\underline{\mathbf{H}}_0$ best. As mentioned above, $\underline{\mathbf{H}}_0$ depicts a circulant matrix defined by $h_0(n)$. As a matter of fact, $h_0(n)$ is an arbitrary discrete function. During the approximation, we attempt to generate shifted copies of $\{h_0(n)\}$ using the two matrices $\underline{\mathbf{D}}$ and $\underline{\mathbf{\Omega}}$. The idea behind this is as follows: In Fig. 5.6(b) and 5.18 examples of the distribution of sensitivity profiles of a sensor array are shown. Naturally, these profiles are displaced from one another according to the orientation of the respective sensors, ρ_l. The key of the new approach is to combine the single profiles in a way that one *joint sensitivity profile* is realized. By shifting $\{h_0(n)\}$ step by step, the joint profile also shifts. A sweep over the entire angular range ($-180° \leq \phi < 180°$) can be performed without altering the shape of the profile. The increment of this sweep is Δr. I.e., the detector appears to be constructed out of N (instead of L) single sensors that are each characterized by the same joint sensitivity profile. Therefore, a *virtually* higher resolution is achieved. The shape of the joint sensitivity profile is defined by the impulse response $h_0(n)$. For this reason, the term *virtual filter interpolation* was chosen. Unfortunately, for non-trivial functions $h_0(n)$, it is impossible to find a set of weighting coefficients so that $\underline{\mathbf{H}}$ becomes *exactly* circulant. Therefore, only an approximation is possible:

$$v(n) \approx v_0(n)$$

Fig. 5.22 shows the joint sensitivity profile approximating $h_0(n)$ at six different shifting angles, using 6, 9 and 25 sensors, respectively[13]. As expected, the approximation quality depends on the number of sensors. For $L = 25$, an excellent result is obtained. The approximation is accomplished by solving a number of linear least square problems. This procedure is discussed in Sec. 5.8.2. For the moment, it is of more interest how the introduced emulated filter can be exploited using discrete signal processing techniques. In the following, the output of the *nominal* filter described by $h_0(n)$ and the corresponding *perfectly* circulant matrix $\underline{\mathbf{H}}_0$, respectively, is denoted $\mathbf{v_0}$ whereas the output of the *approximated* system – defined by $\underline{\mathbf{H}}$ – is $\mathbf{v}$.

We can reconstruct an approximation $\mathbf{s}$ of the radiation pattern $\mathbf{b}$ as illustrated in Fig. 5.23. The following steps are accomplished in this procedure:

Step I. Determine $\mathbf{v}$ from the sensor readings $\mathbf{x}$ and the weighting coefficients $\underline{\mathbf{\Omega}}$, according to Eq. 5.19: $\mathbf{v} = \underline{\mathbf{\Omega}}\,\mathbf{x}$

Step II. Perform a discrete Fourier transform on $\mathbf{v}$, according to Eq. 3:

$$V(k) = \mathsf{DFT}\,[v(n)]$$

Step III. Determine the approximated angular frequency spectrum $\mathbf{S}$:

$$S(k) = \frac{V(k)}{H_0(k)},$$

[13] $h_0(n)$ is defined as depicted in Sec. 5.8.2.

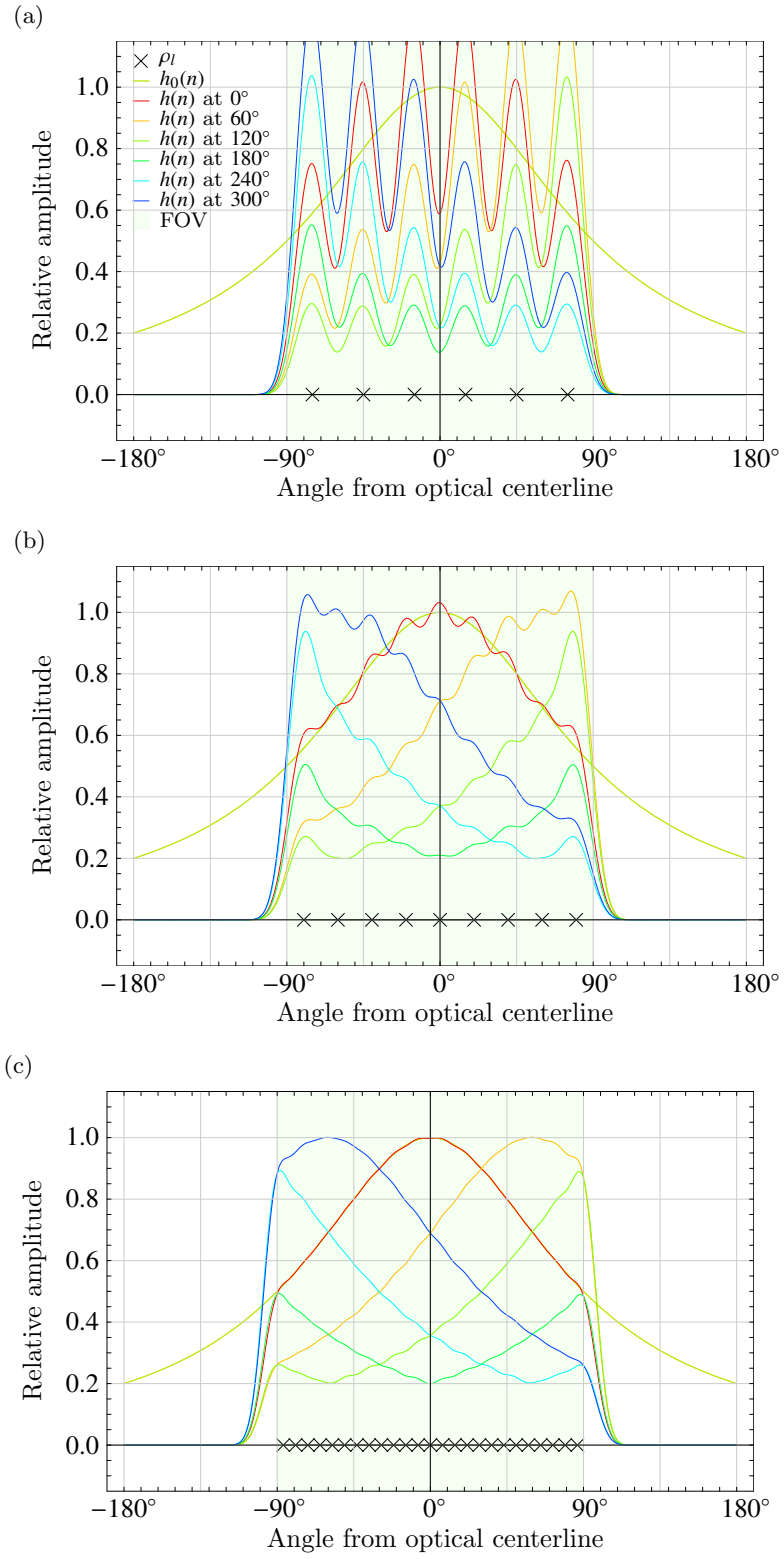


Figure 5.22: Joint sensitivity profiles approximating the impulse response $h_0(n)$ shifted in steps of 60°, with $L = 6$ (a), $L = 9$ (b) and $L = 25$ (c), and $\kappa = 10$, $\tau = 45°$.

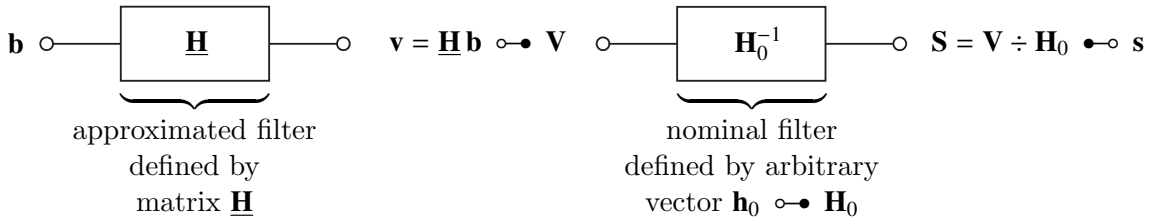


Figure 5.23: Signal processing operations in the LSI system for AOA detection. The operator ÷ indicates element-wise vector division.

with $H_0(k) = \mathrm{DFT}\,[h_0(n)]$. For this operation, $H_0(k)$ must not contain any roots.

Step IV. Reconstruct **b** by performing an inverse discrete Fourier transform on **S**, according to Eq. 4:

$$s(n) = \mathrm{DFT}^{-1}\,[S\,(k)] \approx b(n)$$

Step V. Find the direction α_s of the strongest path by solving $s\,((n - N/2) \cdot \Delta r = \alpha_s) = \max$.

Fig. 5.24 shows the original radiation signal $b(n)$ as well as the reconstructed radiation signal $s(n)$ of a detector array consisting of 15 sensors. As a reference, the result of the low pass interpolation approach, $b'(n)$, is also displayed. A detailed comparison between the low pass and the virtual filter interpolation approaches is presented in Sec. 5.9.

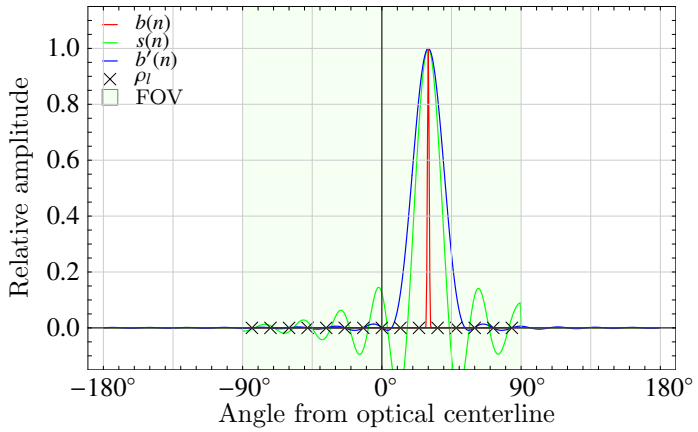


Figure 5.24: Radiation signal recovery using either the virtual filter or low pass interpolation. $\kappa = 10$, $\tau = 90°$, $L = 15$

5.8.2 Determining the weighting matrix

Until now, it has not been mentioned how the weighting matrix $\underline{\mathbf{\Omega}}$ can be generated so that the LSI system introduced in Fig. 5.21 can be realized. As mentioned before, this system is described by its impulse response $h_0(n)$. The cyclic convolution of the radiation sequence $\{b(n)\}$ and the impulse response $h_0(n)$, yielding $\mathbf{v}_0$, can be written as a matrix vector product, with $\underline{\mathbf{H}}_0$ being an $(N \times N)$ circulant matrix defined by $h_0(n)$:

$$\mathbf{v}_0 \quad = \quad \underline{\mathbf{H}}_0 \mathbf{b}$$

Likewise In the following, we *choose* an arbitrary impulse response function $h_0(n)$ whose discrete Fourier transform $H_0(k)$ has no roots in the considered range. E.g. by using the function g_1 with

$$g_1(x) = \frac{1}{1 + (2x)^2}, \quad x \in [-1, 1],$$

$h_0(n)$ is given by

$$h_0(n) = g_1\left(\frac{n}{N/2} - 1\right), \quad 0 \leq n < N. \tag{5.23}$$

Having defined an impulse response, the aim is now to find the best fitting set of weighting factors $\underline{\mathbf{\Omega}}$ such that $\underline{\mathbf{H}}$ becomes circulant as defined in Eq. 5.21. The better the approximation, the better the emulated filter will resemble the filter described by $h_0(n)$. According to Eq. 5.22, perfect approximation is given if

$$\underline{\mathbf{H}}_0 = \underline{\mathbf{\Omega}}\mathbf{D} \quad \text{or} \quad \mathbf{D}^{\mathrm{T}}\underline{\mathbf{\Omega}}^{\mathrm{T}} = \underline{\mathbf{H}}_0^{\mathrm{T}}. \tag{5.24}$$

Since only a part of the detector array is exposed to radiation, we need consider only a subset of the longitudinal range, namely the field of view. Its borders are indexed by $N/2 - F$ and $N/2 + F - 1$, respectively, with

$$F = \left\lfloor \frac{\gamma_{FOV}}{\Delta r} \right\rfloor.$$

Moreover, we introduce two matrices $\underline{\mathbf{A}}$ and $\underline{\mathbf{C}}$, with

$$\underline{\mathbf{A}} = \begin{bmatrix} a(0,0) & a(0,1) & \dots & a(0,L-1) \\ \vdots & \vdots & \ddots & \vdots \\ a(2F-1,0) & a(2F-1,1) & \dots & a(2F-1,L-1) \end{bmatrix}, \quad \text{and}$$

$$\underline{\mathbf{C}} = \begin{bmatrix} c(0,0) & c(0,1) & \dots & c(0,N-1) \\ \vdots & \vdots & \ddots & \vdots \\ c(2F-1,0) & c(2F-1,1) & \dots & c(2F-1,N-1) \end{bmatrix}.$$

$\underline{\mathbf{A}}$ is chosen to be the field of view excerpt of $\underline{\mathbf{D}}^{\mathrm{T}}$. That is, we extract a subset of $\underline{\mathbf{D}}^{\mathrm{T}}$ of size $(2F \times L)$:

$$a(f,l) = d\left(\frac{N}{2} - F + f, l\right), \quad 0 \le f < 2F, \;\; 0 \le l < L$$

Likewise, a $(2F \times N)$ subset of $\underline{\mathbf{H}}_0^{\mathrm{T}}$, denoted $\underline{\mathbf{C}}$, can be defined:

$$c(f,n) = h_0\left(\left(\frac{N}{2} + F - f + n\right) \bmod N\right), \quad 0 \le f < 2F, \;\; 0 \le n < N$$

The reverse indexing of f accounts for the arrangement of impulse response coefficients in the circulant convolution matrix (Eq. 5.21) which is also of reverse order

Eventually, we eliminate the irrelevant (i.e., non-FOV) angular range in Eq. 5.24 by substituting $\underline{\mathbf{D}}^{\mathrm{T}}$ and $\underline{\mathbf{H}}_0^{\mathrm{T}}$ with $\underline{\mathbf{A}}$ and $\underline{\mathbf{C}}$, respectively:

$$\underline{\mathbf{A}}\underline{\boldsymbol{\Omega}}^{\mathrm{T}} = \underline{\mathbf{C}} \tag{5.25}$$

Eq. 5.25 can be split into N linear equation systems

$$\underline{\mathbf{A}}\boldsymbol{\omega}_n = \mathbf{c}_n.$$

with $\mathbf{c}_n$ being the n-th column vector of $\underline{\mathbf{C}}$. Each system is overdetermined because $2F > L$ holds. For a realistic sensor sensitivity pattern, $\underline{\mathbf{A}}$ has full rank. Therefore, a least squares solution exists which can be found by minimizing the Euclidean norm:

$$\|\underline{\mathbf{A}}\boldsymbol{\omega}_n - \mathbf{c}_n\| = \min$$

Solutions to these problems can be found e.g. using the householder method [38]. With each least squares solution $\boldsymbol{\omega}_n$, an approximation to the impulse response of the virtual filter, shifted by n samples (corresponding to an angular shift of $\Delta\phi = n \cdot \Delta r$), is determined.

The question may arise why N equation systems must be solved, although only the field of view, i.e. the range $[N/2 - F, N/2 + F - 1]$ is of interest. Applying the virtual filter, cyclic convolution of sequences of length N is performed. Therefore, weighting coefficients for all angles, not only those of the field of view, must be computed. In order to avoid any "wasted" effort, it is reasonable to redefine N such that it covers only the field of view, i.e. $N = 2F$.

Note that the process of least squares finding need not be executed online, but is a one time task. To save computing and energy resources during operation, it is therefore reasonable to determine the $N \times L$ weighting coefficients $\omega_n(l)$ in advance and store them in non-volatile memory. The footprint of the resulting data set amounts to $N \cdot L \cdot q$ Byte, with q being the number of bytes allocated for one value $\omega_n(l)$. For example, in an arrangement of $L = 15$ sensors, covering a field of view of $2\gamma_{FOV} = 180°$, and using an internal angular range of $N \cdot \Delta r = 360°$ at a resolution of $\Delta r = 0.5°$, and using 32 bit floating point numbers, the data set's memory footprint would be 42.19 kByte. As stated above, the number of equation systems and hence the memory footprint to store the weighting matrix is defined by N and is independent of the field of view. If the sensors have ideal equal sensitivity profiles and are placed symmetrically, and if $\{h_0(n)\}$ is a symmetric sequence then only a fraction of memory is needed.

5.8.3 Computational effort

The overall effort to process one sample set $\mathbf{x}$ is again expressed by the total number of real floating point multiplications and additions. Tab. 5.2 summarizes the number of operations incurred by processing steps I–IV as described in Sec. 5.8.1. Note that the complex divisions in step III can be transformed to complex multiplications because $H_0(k)$ can be computed offline. Hence, $\frac{1}{H_0(k)}$ can likewise be computed offline. As a result, step III can be decomposed to $3N$ real multiplications and $5N$ real additions[14].

Step	*MULTs*	*ADDs*
I.	$N \cdot L$	$(N-1)L$
II.	$\frac{1}{2}N \log_2 N - \frac{3}{2}N + 2$	$\frac{3}{2}N \log_2 N - \frac{5}{2}N + 4$
III.	$3N$	$5N$
IV.	$\frac{1}{2}N \log_2 N - \frac{3}{2}N + 2$	$\frac{3}{2}N \log_2 N - \frac{5}{2}N + 4$
total	$M_N^{\text{filter}} = N \log_2 N + N \cdot L + 4$	$A_N^{\text{filter}} = 3N \log_2 N + N \cdot L - L + 8$

Table 5.2: Computational effort of the filter approach

Note that it was assumed that N is a power of 2 and a split radix FFT algorithm was used. Complex multiplications were split into 3 real multiplications and 3 real additions.

5.9 Comparison of the two approaches

In this section, the two described approaches for signal recovery – low pass and virtual filter interpolation – are compared. In order to motivate the latter method, it has already been shown in Sec. 5.7 that the low pass interpolation approach is susceptible to irregular sensor sensitivity profiles and an unbalanced distribution of sensors. First, the performance of the filter method is evaluated under ideal conditions. Subsequently, we examine the new technique under the condition of an irregular sensor arrangement. For concerns of readability we use the shortcuts LPI (low pass interpolation) and VFI (virtual filter interpolation) in the following. Furthermore, according to Eq. 5.12, we define the absolute AOA error $\epsilon_s(\alpha)$ of the measured angle α_s:

$$\epsilon_s(\alpha) = |\,\alpha - \alpha_s\,|$$

Fig. 5.25 shows the mean measurement error using the LPI method $\langle \epsilon_{b'}(L) \rangle$ and the VFI approach, $\langle \epsilon_s(L) \rangle$, respectively, for differently shaped sensitivity profiles. In both cases, the low pass interpolation method performs significantly better. The gain depends on the sensitivity profile of the sensors. This yields the conclusion that for perfectly manufactured and mounted sensors that are arranged in an equidistant manner, the LPI method beats the VFI approach.

[14] See note on decomposition of multiplications of two complex numbers in Sec. 5.6.3.

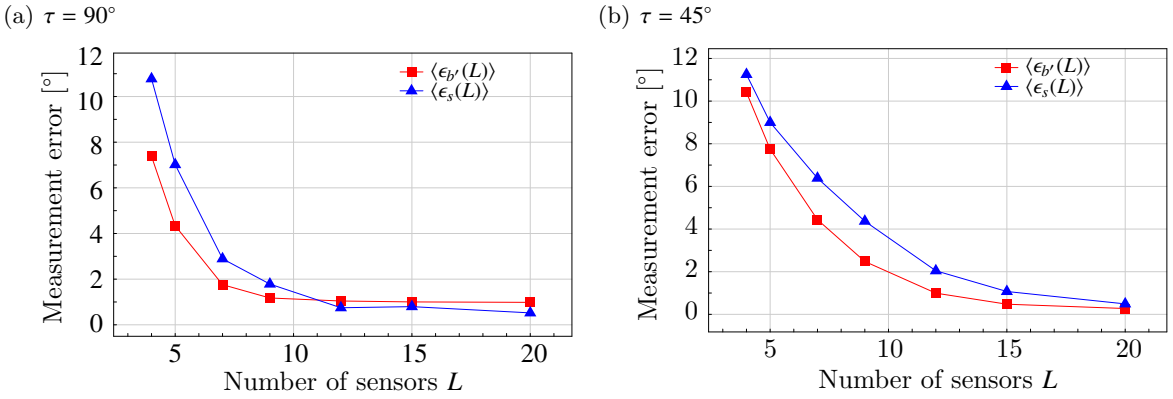


Figure 5.25: Low pass vs. virtual filter using a regular array $\kappa_0 = 10$, $\sigma_\kappa = 0$, $\sigma_\tau = 0$, $\sigma_\rho = 0$, $\sigma_a = 0$

The performance of the virtual filter method in an unbalanced detector was subject to the next series of investigations. An excerpt of the results can be seen in Fig. 5.26. The absolute measurement errors of the two methods are even-tampered for nine sensors, whereas for twelve sensors, the filter method clearly outperforms the low pass approach. Also, contrary to the LPI method, the filter approach lacks any significant accuracy outliers. This fact is an important benefit of the new method.

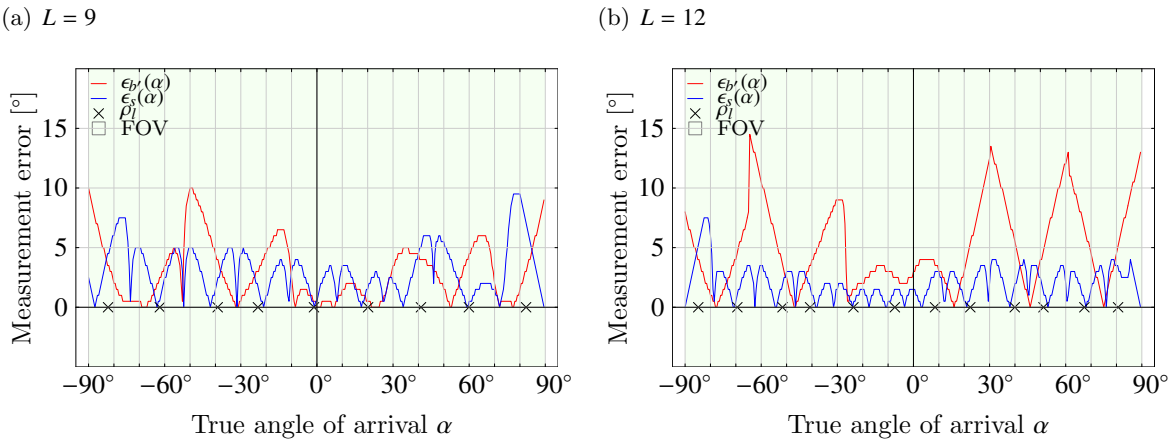


Figure 5.26: Low pass vs. virtual filter approach using an irregular array, $\kappa_0 = 10$, $\sigma_\kappa = 0$, $\tau_0 = 90°$, $\sigma_\tau = 0$, $\sigma_\rho = 2°$, $\sigma_a = 0.1$

The mean measurement error against the number of sensors of an unbalanced detector array is displayed in Fig. 5.27. At a low number of sensors, the low pass interpolation method yields better results. At $L = 9$ (a) and $L = 12$ (b), respectively, both approaches perform equally well, and beyond this number of sensors, the filter method is the superior technique. Interestingly, only the VFI approach, $\langle \epsilon_s(L) \rangle$, yields a monotonic decreasing graph. Contrary

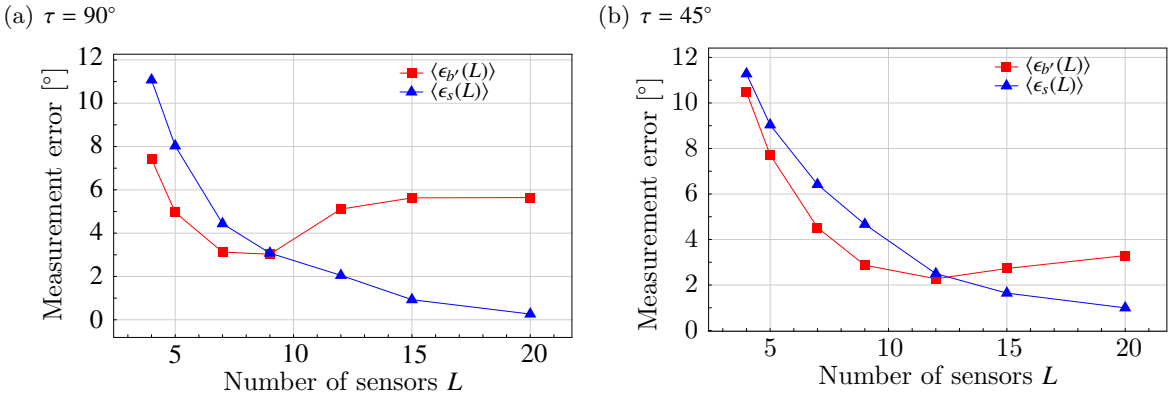


Figure 5.27: Low pass vs. virtual filter approach using an irregular array, $\kappa_0 = 10$, $\sigma_\kappa = 0$, $\sigma_\tau = 0$, $\sigma_\rho = 2°$, $\sigma_a = 0.1$

to this, the graph of the LPI method, $\langle \epsilon_{b'}(L) \rangle$, decreases for a low and rises for a high number of sensors. E.g., in Fig. 5.27(a), the detector seems to perform better using 7 than using 20 sensors. The reason for this behavior can be explained as follows: the angular distribution error, defined by σ_ρ, is constant. The nominal angular displacement between two sensors, i.e. the sampling period Φ, is large for a low number of sensors but decreases as further sensors are added. Therefore, the rate σ_ρ/Φ increases as L increases, i.e. the impact of σ_ρ increases.

5.9.1 Computational effort

The arithmetic effort for both approaches is covered in Sec. 5.6.3 and 5.8.3, respectively. Since the floating point multiplication is a significantly more complex operation than the addition, only the number of multiplications is considered to compare the two approaches. Low pass interpolation using direct convolution – provided a number of simplifications – yields $M_N^{\text{conv direct}} = (N/2+1) \cdot L$ multiplications (Eq. 5.13), whereas the virtual filter approach requires $M_N^{\text{filter}} = N \log_2 N + N \cdot L + 4$ (see Tab. 5.2). The complexity of both methods depends on L. Regardless of the number of sensors, the low pass approach is significantly faster ($M_N^{\text{conv direct}} < M_N^{\text{filter}}$). In an experimental implementation using an ATmega 8-bit microcontroller at 16 MHz the speed of both methods has been evaluated. The execution time to determine the angle of arrival vs. the sequence length N is shown in Fig. 5.28. $T_N^{\text{conv direct}}(L = 10)$ and $T_N^{\text{filter}}(L = 10)$ denote the execution time for the low pass interpolation and the filter approach, respectively. At $\log_2 N = 7$, i.e. $N = 128$ the result of both approaches is obtained after less than 200 ms. Therefore, the update rate R of the detector is in this case $R \geq 5/\,\text{s}$. This is sufficient for most indoor applications. Yet, the limited sequence length of $N = 128$ might be insufficient: given a resolution $\Delta r = 0.5°$, an angular range of only $\pm 64°$ can be covered. If Δr is increased, a larger range can be observed, but the resolution might become unacceptable to certain applications. On the other hand, increasing the sequence length to e.g. $N = 512$ raises the execution time – and thus, the detector update rate – to

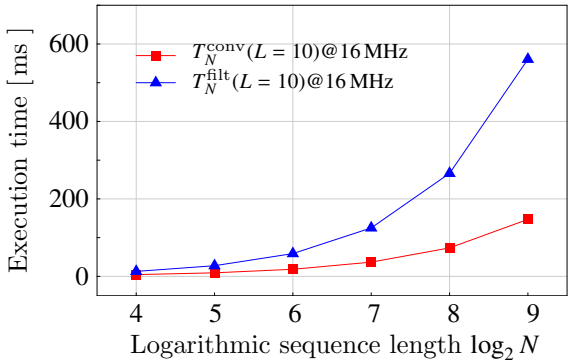


Figure 5.28: Computational effort of the two approaches using an ATmega 8bit microcontroller

a possibly unreasonable level. It must also be remembered that power consumption is an important implementation criterion. Since it rises with execution time, it constitutes another reason why the length of N must not exceed a certain level.

However, the implementation is a straight-forward C implementation using 32-bit floating point numerics. A much faster code with reasonable accuracy can be obtained using fixed point arithmetics, and possibly assembler language. Nevertheless, the example implementation shows that a very low power CPU can handle the algorithmic effort in reasonable execution time.

5.10 Multiple targets

To this point we have only discussed the presence of *one* source of radiation. If more sources are available, different solutions to discriminate them are at hand:

- The straight-forward approach to allow for multiple targets is to reserve time slots for each target. Only in its own time slot, a target may emit radiation. I.e., concurrent emission of radiation by multiple mobile targets is prevented. As a result, only one peak of radiation is determined by the detector array. This approach requires an arbitration of the optical channel. For this purpose, all targets must be equipped with communication modules.
- If channel arbitration is an unacceptable solution (possibly for concerns of increased cost), an alternative approach must be found, which allows for the concurrent emission of radiation by multiple targets. In this case, the reconstructed radiation profile may have several concurrent peaks. Each peak must be assigned to one particular source since otherwise, no triangulation can be performed. For this purpose, each target must emit a unique code signature. By means of cross correlation we can then determine whether or not the detector has received a sequence from a particular target.

5.11 Detector precision in the presence of noise

Until now, we have ignored any noise generated in the detector elements. In this section, we analyze the detector's performance at different signal to noise ratios. As discussed in Sec. 5.3, we can separate the generated noise into external and internal components – ambient noise produced by background lighting, as well as receiver noise produced by front-end electronics. For a well designed preamplifier, ambient noise is dominant over receiver noise. We can consider the noise as white and Gaussian.

Before any investigation on the impact of noise on the measurement system can be made, we must define a measure to relate the received signal power with the noise power. When doing this we must keep in mind that we deal with an *array* of sensors each of which has a specific directivity profile. The total instantaneous received signal power of this array is given by

$$P_{tot} = \sum_{l=0}^{L-1} {x_l}^2. \tag{5.26}$$

Until now, we have only taken one snapshot of concurrent sensor readings $\{x_l\}$ into account. When we carry out G experiments we obtain a set of sequences whose elements are given by $x_l(g)$, with $0 \leq g < G$. The average power is given by the mean of the total power of all G experiments:

$$P_{avr} = \frac{1}{G} \sum_{g=0}^{G-1} P_{tot}(g) = \frac{1}{G} \sum_{g=0}^{G-1} \sum_{l=0}^{L-1} {x_l}^2(g) \tag{5.27}$$

We assume that the obtained signal power is equal in all experiments. The signal to noise ratio is commonly defined by the ratio of signal power to noise power, i.e.

$$SNR = \frac{P_{avr}}{{\sigma_N}^2}.$$

${\sigma_N}^2$ is the average noise power. We assume that statistically the same level of noise is generated in each sensor. Hence the average power contributed by sensor l, ${\sigma_{N,l}}^2$, is

$${\sigma_{N,l}}^2 = \frac{{\sigma_N}^2}{L} = \frac{P_{avr}}{SNR \cdot L}.$$

We can now generate Gaussian distributed noise samples for each sensor l with a standard deviation

$$\sigma_{N,l} = \sqrt{\frac{P_{avr}}{SNR \cdot L}}.$$

The impact of noise was simulated using a detector array consisting of 9 and 12 sensor elements, respectively. The aberrations statistics of this array were chosen as in the previous simulations. The measurement error against the SNR, using the LPI and VFI approach, respectively, is shown in Fig. 5.29. For each SNR value, $G = 1000$ sample sets of sensor readings were taken and averaged. Subsequently, for each SNR value, the AOA error was averaged over

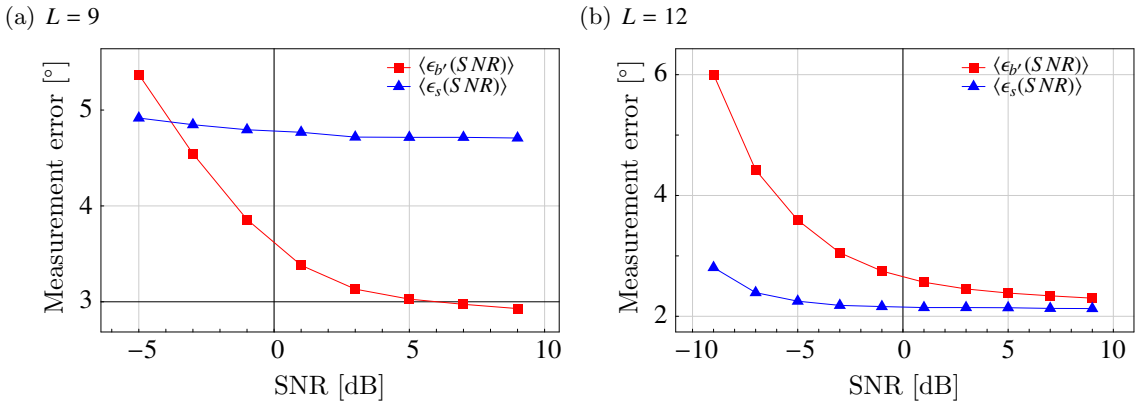


Figure 5.29: Average AOA measurement error vs. signal to noise ratio. $\kappa_0 = 10$, $\sigma_\kappa = 0$, $\tau_0 = 45°$, $\sigma_\tau = 0$, $\sigma_\rho = 2°$, $\sigma_a = 0.1$, $G = 1000$

the entire field of view, resulting in the error $\langle \epsilon_{b'}(SNR) \rangle$ and $\langle \epsilon_s(SNR) \rangle$, respectively. In all simulations, the sensor aberrations were kept constant. Both plots show that the error of both approaches converges to a theoretical minimum which is defined by the physical constraints of the detector array (Number of sensors, directivity characteristics, etc.). These minimum values can also be found in the results shown in Fig. 5.27(b)[15]. Obviously, the VFI is more robust against noise: For a decreasing SNR, the error of the low pass approach, $\langle \epsilon_{b'}(SNR) \rangle$, increases much faster than the error of the virtual filter approach, $\langle \epsilon_s(SNR) \rangle$. Therefore, although it might not be the better solution for *ideal* sensor arrays, the VFI approach may outperform the low pass method for a low SNR. For heavily aberrated sensitivity profiles, VFI is the superior approach at all times.

5.12 Simulation of the indoor IR channel

In the previous sections we have assumed that radiation is only received from directions in the horizontal (x-y) plane. This fact was assured by the transmitter design presented in Sec. 5.2.2. For other transmitters, we must take other directions of incidence into account whose z-component is unequal to zero (e.g., a ray that is reflected at the floor). In order to examine the impact of reflections in a typical room, an infrared raytracing engine was used which was developed at the University of Paderborn, Germany [48]. Fig. 5.30 shows an image generated using this engine.

The image shows the near infrared radiation measured by a virtual spherical camera, whose

[15]The results of Fig. 5.27(b) and 5.29(b) do not *completely* match since a different set of sensor aberrations was generated. However, the general conclusion we can draw is not affected by this fact.

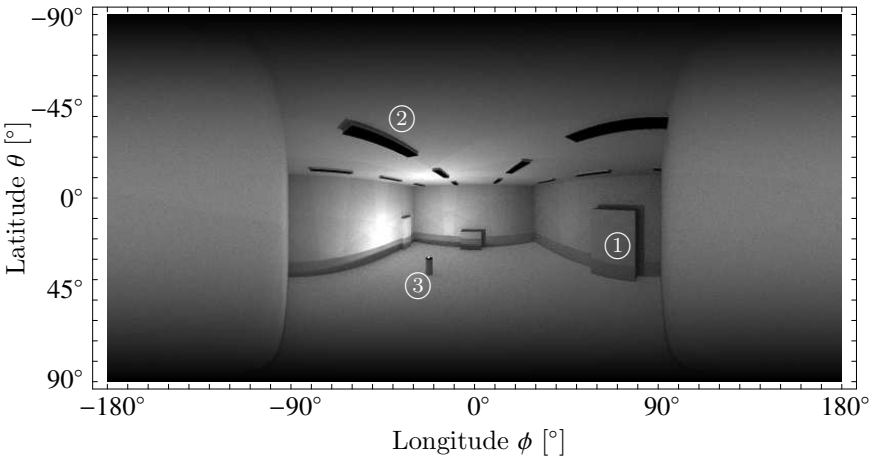


Figure 5.30: Scene acquired by a virtual spherical camera

output is projected into the image plane. The axes represent the longitudinal and latitudinal ranges, respectively. The scene modeled for this image consists of a room of the dimensions $9 \times 10 \times 5\,\mathrm{m}^3$ containing a closet, two cabinets (①) and a few ceiling lights (②). Placed on a solid cylinder of 1 m height (③) is an isotropic radiator with the shape of a sphere. Its lower hemisphere resides within the cylinder. Therefore all radiation in a direction below the equator is absorbed. The surface properties of walls, floor, ceiling and objects were chosen to match the respective physical properties of the real world. For instance, the degree of specular and diffuse reflection at walls corresponds to typical rough plaster.

In the following, we replace the virtual sphere camera by the detector array. Note that the detector is unchanged, i.e. the sensor elements are still oriented in the plane. Only the inclination, denoted η, is introduced. The inclination is illustrated in Fig. 5.31.

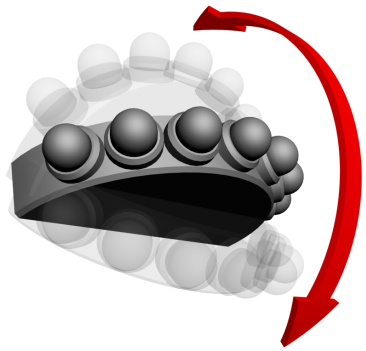


Figure 5.31: Inclination of the detector array

Furthermore, the radiation profile $\mathbf{b}$ is now substituted by a $(M \times N)$ matrix $\mathbf{B}$ which

corresponds to the $(N \times M)$ bitmap[16] shown in Fig. 5.30. This radiation matrix is expressed by

$$\underline{\mathbf{B}} = \underbrace{\left.\begin{bmatrix} b(0,0) & b(0,1) & \dots & b(0,N-1) \\ b(1,0) & b(1,1) & \dots & b(1,N-1) \\ \vdots & \vdots & \ddots & \vdots \\ b(M-1,0) & b(M-1,1) & \dots & b(M-1,N-1) \end{bmatrix}\right\} M \text{ latitude angles}}_{N \text{ longitude angles}}$$

with $M = \left\lfloor \frac{180^\circ}{\Delta r} \right\rfloor$ and $N = \left\lfloor \frac{360^\circ}{\Delta r} \right\rfloor$.

Since radiation is received from all directions, we must also take into account that the sensor elements' sensitivity characteristic is of uniaxial symmetry (see Fig. 5.32). A sensor

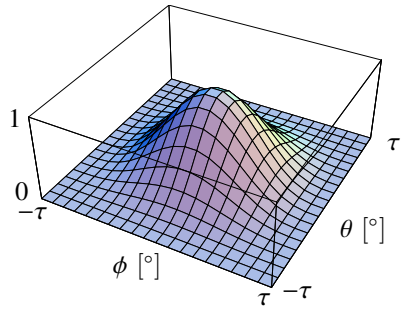


Figure 5.32: Sensor sensitivity vs. angular displacement ϕ and θ ($\kappa = 3$, $\tau = 75^\circ$)

reading x_l can now be determined in the following way: The portion of radiation "seen" by a single sensor l can be expressed by the product of the sensor characteristic (Fig. 5.32) and a certain square subset of the radiation matrix. This procedure is graphically shown in Fig. 5.33. The square selection is chosen according to the angular displacement of the sensor, ρ_l, and the inclination η of the sensor array.

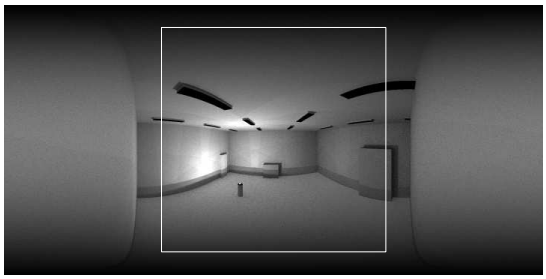
Mathematically, we can transform the *two dimensional* radiation bitmap $\underline{\mathbf{B}}$ into the *one dimensional* radiation profile $\mathbf{b}$. Subsequently, we can apply one of the interpolation approaches introduced in the previous sections. The transformation process is explained in the following.

$\underline{\mathbf{B}}$ can be transformed to $\mathbf{b}$ under consideration of the latitudinal sensitivity characteristic of the sensor elements. Similar to the longitudinal characteristic $d_l(n)$ given in Eq. 5.7, we define the latitudinal characteristic $\mathbf{d'} = \{d'(m)\}^{\mathrm{T}}$ with

$$d'(m) = \begin{cases} cos^{\kappa}\left((m \cdot \Delta r - \eta - 90^\circ) \cdot \frac{90^\circ}{\tau} \cdot \frac{\pi}{180^\circ}\right), & \text{if } -\tau \le m \cdot \Delta r - \eta - 90^\circ < \tau \\ 0, & \text{otherwise,} \end{cases}$$

[16] The virtual camera has $N \times M$ pixels.

(a) Subset of the radiation bitmap incident on a single sensor l located at $\rho_l = 0°$, indicated by a square



(b) Resulting radiation as 'seen' by the sensor

Figure 5.33: Radiation perceived by an individual sensor. The sensor array is orientated in the horizontal plane, i.e. it is not inclined ($\eta = 0°$)

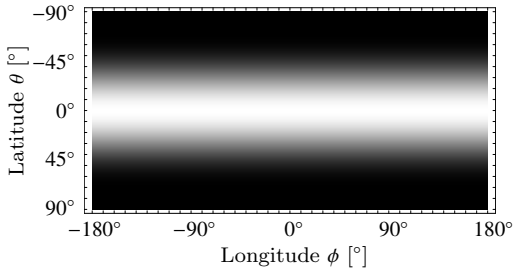
where $m = \left\lfloor \frac{90°+\theta}{\Delta r} \right\rfloor \bmod M$. Each column of $\underline{\mathbf{B}}$ is element-wise multiplied by the latitudinal sensitivity profile $\mathbf{d}'$. Then the radiation perceived over the entire latitudinal range is summed up. In matrix notation, these operations summarize to

$$\mathbf{b} = \underline{\mathbf{B}}^{\mathrm{T}}\mathbf{d}'.$$

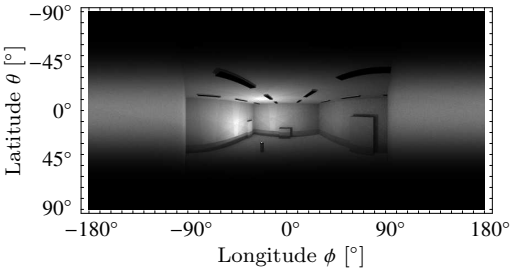
The transformation is illustrated in Fig. 5.34. It can be seen that under realistic conditions, the resulting signal $b(n)$ (Fig. 5.34(c)) has a very narrow and dominant peak at the direct line of sight between the radiation source and the detector array[17]. We can conclude that the degree of specular reflection of near infrared radiation is very low for the materials used in the simulation. I.e., the direct echos that can be observed have a very low intensity. This result complies with real world measurements described in the literature.

[17]Note that the intensity scale in Fig. 5.34(b) is clipped. As a result, the direct line signal (little white spot on top of the cylinder) appears to have the same intensity as the direct echo paths (the large bright spots at walls and ceiling). As a matter of fact, the direct line signal is many times larger than any echo signal. The intensity scale for this figure was clipped in order to demonstrate the general idea of the transformation from $\underline{\mathbf{B}}$ to $\mathbf{b}$.

(a) Relative latitudinal sensitivity



(b) Latitudinal sensitivity multiplied by radiation matrix $\underline{\mathbf{B}}$



(c) Resulting radiation profile $\mathbf{b}$

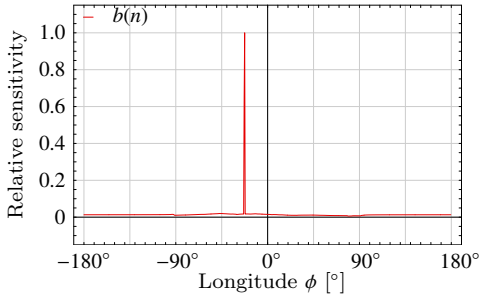


Figure 5.34: Transformation from $\underline{\mathbf{B}}$ to $\mathbf{b}$. $\kappa = 3$, $\tau = 75°$, $\eta = 0°$

Chapter 6

Summary and outlook

In this work, several aspects of indoor localization have been discussed, and new approaches proposed to allow for the use of low cost components. Providing physical position information of vehicles, humans and objects is one of the key issues in robotics and pervasive computing: It enables uncounted applications in the industrial, commercial, medical, military, and home environment. Since its existence, mankind has dealt with the problem of localization, and hence we are familiar with many of these applications. However, new location and context aware applications are emerging, demanding increased localization accuracy at low cost. A lot of these future applications are situated indoors.

In Ch. 2 of this thesis, common absolute localization techniques have been discussed: triangulation, trilateration, and a variant of the latter, hyperbolic localization. In either method, fixed reference units are deployed in the environment. These reference units interact with the mobile target whose position is to be determined by transmitting and receiving signals. This signal transmission enables the measurement of angles, distances, or distance differences among the devices – depending on the applied localization technique. From the results obtained, the position of the mobile target can be determined. For this purpose, the geometric arrangement is exploited. Common ways of range and angle measurements have been described. The most relevant implementations of indoor localization systems have been briefly discussed and compared in Ch. 3.

The measurements mentioned above pose the most fundamental challenges on a localization system, especially if its components must be inexpensive. Therefore two particular problems were selected in this work: time measurement and synchronization in hyperbolic localization, as well as angle measurement in triangulation.

6.1 Low cost time measurement and synchronization in hyperbolic localization

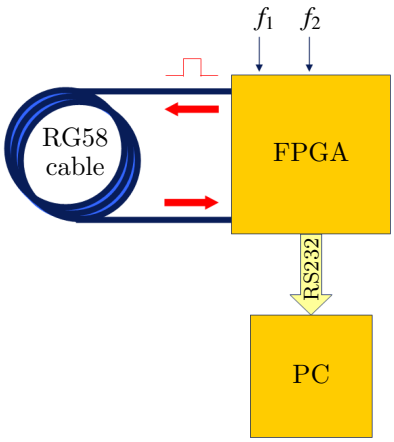
A common approach for location estimation is hyperbolic localization. Following this method, time difference of arrival (TDOA) measurements among a mobile and several fixed reference units are performed. For this purpose, these units are equipped with clocks which must be synchronized among each other. This is particularly critical when electromagnetic waves are used which propagate at the speed of light. In this case clocks must operate at a very high precision and are therefore complex and expensive. As a consequence, current implementations require costly synchronization hardware, or they avoid high accuracy synchronization by using hybrid signal technologies like RF and ultrasound.

In Ch. 4 of this work, a time acquisition approach for indoor hyperbolic localization has been presented that allows for the use of electromagnetic waves. The method has two major advantages: a) A high location accuracy is obtained in spite of using relatively low frequency and unstable or drifting clock oscillators. This is achieved by transmitting bursts of pulses (instead of a single pulse) following a certain transmission pattern, and subsequent time averaging. As a consequence, the localization system can be implemented in inexpensive standard CMOS without the need for sophisticated high speed circuitry. b) The participating units' clocks are implicitly synchronized which further reduces system complexity and cost. Inaccuracies of the oscillators are combated in the following way: the impact of short term instabilities causing jitter is delimited by dividing the measurement procedure into smaller burst slots. As a result, the system is more frequently resynchronized and hence jitter cannot grow unboundedly. Long term clock drift due to aging is compensated by adding an additional reference unit and carrying out clock bias compensation measurements. Simulation results have been provided that indicate how the achieved accuracy depends on various factors such as burst length and jitter. Oscillators in the order of 100 MHz can usually resolve only 3 m. With the approach proposed here, a significantly higher accuracy can be obtained. In an example configuration using 12 ps RMS clock jitter at all units, the standard deviation of the measurement error could be decreased to below 4 cm for measurement times of 30 ms and above. The approach was shown for only one mobile target. For a setting of multiple mobile units, two solutions come at hand:

- Beacon signals contain a digital identifier. The mobile units blindly emit bursts of beacon signals in a random pattern. It is accepted that in some cases, mobile units may transmit at the same time such that their beacon signals interfere. Signal collisions can be identified and the bad signals discarded. Subsequently, signal bursts must be remitted.

- A global arbiter reserves time slots for the localization of each mobile target. Of course, this implies some effort: Mobile units within range of the arbiter must announce themselves and apply for a time slot. However, in this case no digital identifier is required because normally, no collisions may occur.

(a)



(b) Implementation using a Xilinx Virtex II FPGA

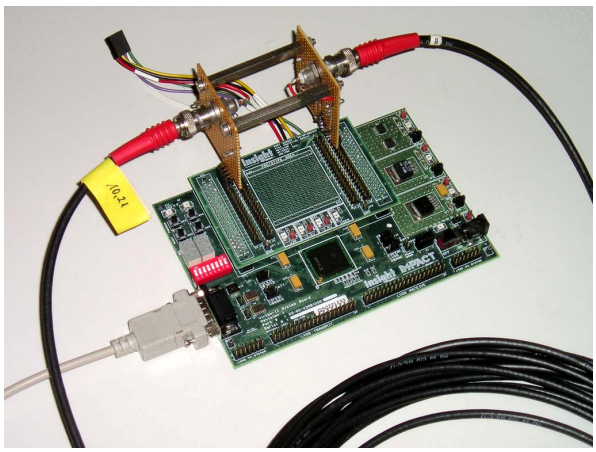


Figure 6.1: Experimental setup for cable length measurements, based on the proposed measurement approach

Another modification of the system that comes to mind is the deployment of additional reference units, in order to e.g. overcome bad measurements of single reference units. Different solutions have been proposed for a redundant hyperbolic localization system, e.g. maximum likelihood estimation [21]. However, the proposed measurement approach needs some modifications in this case. The reference measurements for clock frequency bias and clock synchronization must be extended to synchronize all reference units. The localization measurement (where the mobile unit acts as transmitter) can remain unchanged – except for the fact that now multiple TDOAs are obtained by different pairs of receivers. The modifications in the measurement procedure are straight forward. However, the more reference units are involved, the lower the resynchronization rate of a single reference unit. Hence, the localization precision may decrease if too many redundant reference units are deployed.

6.2 Alternative application using the proposed time measurement approach

The time measurement approach proposed in this thesis can be applied to alternative problems. E.g., the length of untwisted cables can be determined by measuring roundtrip times. This application was evaluated in an experimental setup using a Xilinx Virtex II FPGA, as shown in Fig. 6.1. The core system consists of the following functional blocks: transmitter, receiver, preprocessor and communication unit. The IC is clocked by two crystal oscillators ($f_1 = 100\,\text{MHz}$ and $f_2 = 96\,\text{MHz}$). Results are presented using a PC connected via RS232. No external circuitry is needed except for a terminating resistor that avoids signal reflections within the cable. The transmitter continuously emits electric pulses which are after a *varying* time detected by the receiver unit. This variance in detection time is a result of using asyn-

chronous oscillator clocks for sender and receiver, as explained in Ch. 4. The roundtrip time is measured by a counter running at a frequency of the sender's clock, f_1. The preprocessor unit accumulates the measured time for a given number of estimations before transmitting the result over the RS232 link. This helps to overcome the low data rate of this connection. After 16000 estimation cycles a precision of ±0.5% was achieved. I.e., for 1 m and 20 m cables we measured 100 ± 0.5 cm and 20 ± 0.1 m, respectively. The total estimation time was 5 ms. The rising inaccuracy for longer cables is caused by path loss in the cable. The method is highly susceptible to aberrations of the cable properties. E.g., cables of the same length and compliant to the same norm, but manufactured by different companies, yielded reproducible different results.

6.3 Angle of arrival measurement using an infrared detector array

In Ch. 5 of this work, angle measurement of incident infrared radiation using a detector array has been discussed. When at least two such detector arrays are deployed at walls, and the mobile unit emits an infrared signal into the horizontal plane, then its position can be determined by triangulation. Common designs of angle diversity receivers have been briefly described. They consist of multiple sensing elements whose angular sensitivity characteristics point into different directions, spanning dedicated sectors. While in communications, the sectors are aimed to have sharp edges, in the angle of arrival (AOA) approach presented here, the sensitivity profiles of adjacent sensor elements may overlap. Commonly it is presumed that all sensor elements incorporate equally shaped angular sensitivity functions. In this case the detector array can be treated as a sampling system with a sampling period defined by the angular displacement of two adjacent sensor elements. However, this is usually not the case, particularly when the detector is aimed to be low cost. Sensor aberrations originate from attached optical components such as lens and filter, as well as manufacturing and assembly inaccuracies. For example, inexpensive arrays are equipped with a single silicon imaging lens which distorts the scene on the sensor area. As a result, the sampling model will fail for real world detectors. A novel method named *virtual filter interpolation* (VFI) has been presented that incorporates the individual angular sensitivities of the sensor elements using a linear least squares approach. The new approach has been discussed and compared with the ideal sampling method. While the latter performs better for an ideal detector array, it is clearly outperformed by the VFI method when irregularities of real world detectors are taken into account. The VFI approach has a higher computational complexity, though. However, it has been shown that even an 8-bit microcontroller can perform an AOA measurement in far less than a second, at a reasonable resolution.

The direction finding approach presented in this chapter can be applied to a number of applications. Even the technology may vary. As an example application, mobile robot localization was picked. Robots mounted on wheels do usually not jiggle as much as human bodies when moving. Therefore, the transmitter introduced in Sec. 5.2.2 is well suited for

robot localization, but inappropriate to be worn by humans.

6.4 Outlook: Localization of humans using thermoelectric detectors

The detector design and the discussed direction finding approaches can nevertheless be used to localize humans and other hot blooded creatures. For this purpose, a different detector technology must be chosen that senses the infrared radiation naturally emitted by humans. A source of radiation can be detected as long as the ambient temperature differs by a certain amount from the human skin temperature. The peak wavelength of the natural IR emission by humans is approx. 9.4 µm. This figure is obtained using Wien's displacement law that states that there is an inverse relationship between the peak emission wavelength λ_{max} of a black body[1] and its temperature T,

$$\lambda_{max} = \frac{2898\,\mu\text{m} \cdot \text{°C}}{T}.$$

The human skin temperature is usually around 34 °C. Although the human body is not truly a black body, it is a good approximation[2]. Photodiodes discussed in the previous sections are inappropriate for the peak wavelength of 9.4 µm: For rising wavelengths, the effort to cool the photodetector and thus reduce intrinsic noise also rises. Contrary to photodiodes, many *thermal detectors* are not afflicted with this disadvantage. However, they have generally a lower detectivity.

A well known thermal detector is the pyroelectric sensor which is widely used for motion sensing in surveillance and intruder detection applications. Hence these devices are very low cost. Unfortunately, pyroelectric sensors are sensitive only to *fluctuations* of infrared radiation. In order to generate these fluctuations, the monitored area is divided into alternating active and blind zones. This is achieved by a specifically designed multi-segmented lens. When the a human body moves between these zones the sensor interprets this movement as occupancy. Unfortunately, no localization or tracking of moving persons is possible because the occupied zones cannot be identified. That is, the sensor can only tell *that* a source of IR radiation is moving in the field of view, but not *where*. Static or slowly human bodies cannot be detected at all. Because of the attached lens, the sensor has an almost constant sensitivity in all directions, thus no angular characteristic can be exploited.

For localization of humans exploiting their natural emission of IR radiation, other thermal detectors such as *thermopiles* can be used [35, 89]. These devices are thermoelectric sensors based on the Seebeck effect. They consist of a pile of thermocouples electrically connected in series in order to increase the output voltage. A thermocouple is usually formed by a dissimilar metal junction. Hence, the series of thermocouples results in a chain of junctions between

[1] A black body is a theoretical model of an object that absorbs all incident electromagnetic radiation. It emits radiation whose spectral intensity depends on its temperature.

[2] The emissivity of a black body is by definition $\varepsilon = 1$, the emissivity of the human body is $\varepsilon = 0.98$ [78].

two metals a and b. All junctions $a \rightarrow b$ – usually referred to as *hot junctions* – are thermally coupled with the absorber membrane. The junctions $b \rightarrow a$ (*cold junctions*) are coupled to the device frame, which is a large integrated or external heat sink. The amount of infrared radiation received by the membrane is proportional to the electrical potential obtained. Accordingly, the resulting voltage of the thermopile is proportional to the temperature difference of the hot and cold junctions, i.e. the membrane and the heat sink. For absolute temperature measurement, a reference sensor (typically a thermistor) is integrated into the frame. If the heat sink is sufficiently large, these devices can be operated uncooled.

Integrated thermopiles have become inexpensive in recent years (although they are still significantly more expensive than standard photodiodes). They are widely used in non-contact temperature measurement in medical, automation and control applications. Integrated thermopiles are available as single element devices and arrays (e.g., 1×8 line arrays or 4×4 matrix arrays).

If multiple targets (e.g., humans) are located in the monitored area, the following problem may arise: The reconstructed radiation profile may contain multiple concurrent peaks. Unfortunately, the naturally emitted radiation of humans is static[3]. Therefore, neither of the proposals given in Sec. 5.10 can be realized: No channel arbitration and no modulation of the emitted radiation by a unique signature is possible. As an inevitable side effect, no identification of humans can be carried out. Therefore, only stochastic methodologies remain available in order to track people.

[3]at least for the sensors and application discussed here

Appendix A

Illustration of hyperbolic localization

In this section, a graphical analysis of hyperbolic localization is presented. For this purpose each time difference of arrival measurement is treated individually. An equation is developed that yields all possible values x_m and y_m with respect to a given (i.e., measured) value of d_1. This procedure is repeated regarding a given value of d_2. Subsequently the resulting graphs are plotted and their intersections discussed.

We compile the equation for a transformed arrangement of reference units, where a and b are rotated by ϕ and translated by Δx, Δy such that $a' = (0, 0)$ and $b' = (l_{ab}, 0)$ (see Fig. A.1) with

$$\phi = \arctan \frac{y_b - y_a}{x_b - x_a}$$

and

$$\Delta x = x_a, \quad \Delta y = y_a.$$

With

$$l_{ab} = \sqrt{(x_b - x_a)^2 + (y_b - y_a)^2}$$

the following equation system for $(x_{m'}, y_{m'})$ can be set up against parameter α:

$$\begin{aligned} y_{m'} &= \frac{\sin \alpha}{2} \cdot \frac{d_1^2 - l_{ab}^2}{d_1 - l_{ab} \cos \alpha} \\ x_{m'} &= \frac{y_{m'}}{\tan \alpha} \end{aligned}$$

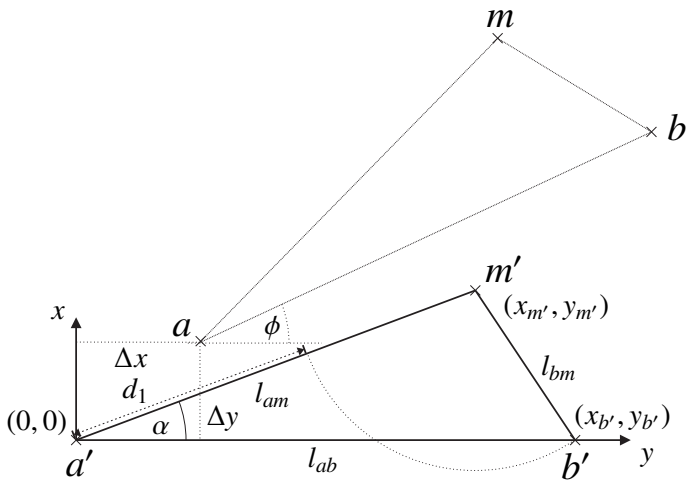


Figure A.1: Geometric arrangement

Subsequently, x_m and y_m can be computed from $x_{m'}$ and $y_{m'}$ by inverse coordinate system transformation. The resulting equation system is

$$\begin{aligned} x_m &= \frac{1}{2} \cdot \frac{d_1^2 - l_{ab}^2}{d_1 - l_{ab}\cos\alpha}(\cos\alpha\cos\phi - \sin\alpha\sin\phi) + \Delta x \\ y_m &= \frac{1}{2} \cdot \frac{d_1^2 - l_{ab}^2}{d_1 - l_{ab}\cos\alpha}(\cos\alpha\sin\phi + \sin\alpha\cos\phi) + \Delta y \end{aligned}$$

The graph of this system will approximate its asymptotes at $\pm\alpha_{max}$ where

$$\alpha_{max} = \arctan\frac{\sqrt{l_{ab}^2 - d_1^2}}{d_1}.$$

With a parameter range of $[-\alpha_{max} \ldots \alpha_{max}]$ the graph consisting of all points (x_m, y_m) yielded by d_1 can be plotted. A similar equation system can be found for d_2, i.e. when using reference units a and c.

Fig. A.2 shows three instances of each of the two graphs, with different values for d_1 and d_2. The points of intersection of these graphs indicate the final position (x_m, y_m) obtained by the respective pair of d_1 and d_2 values.

The resulting equations yield hyperbolas. For this reason, the positioning method is commonly referred to as *hyperbolic localization*. As can be seen, three different situations may occur: the graphs intersect either once, twice or not at all. Two intersection points indicate an ambiguousness for this particular case, i.e. the position cannot unambiguously be determined (bright area in Fig. 4.2). A single point yields one single position (dark area in Fig. 4.2). No intersection indicates that no position can be computed from the given set of distance differences, perhaps due to an erroneous measurement.

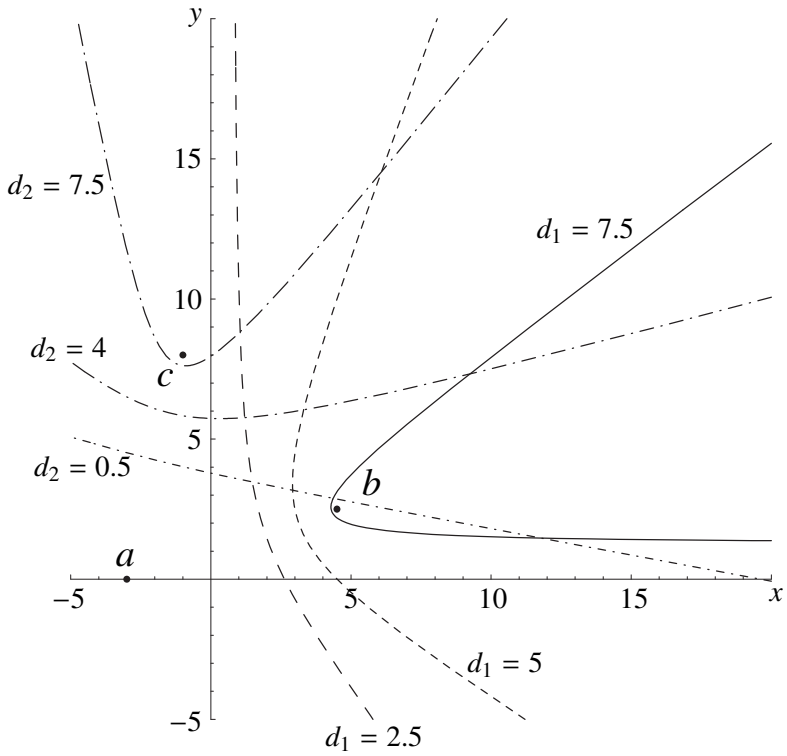


Figure A.2: Intersection of hyperbolas plotted for various positive d_1 and d_2 values. This plot shows the three situations that may occur: a) One point of intersection (e.g. $d_1 = 5, d_2 = 4$), b) Two points of intersection (e.g. $d_1 = 7.5, d_2 = 0.5$), c) No intersection ($d_1 = 7.5, d_2 = 7.5$).

Appendix B

Closed-form solution for hyperbolic localization

The position of the mobile node can be obtained by straight forward evaluation of the geometric arrangement shown in Fig. B.1.

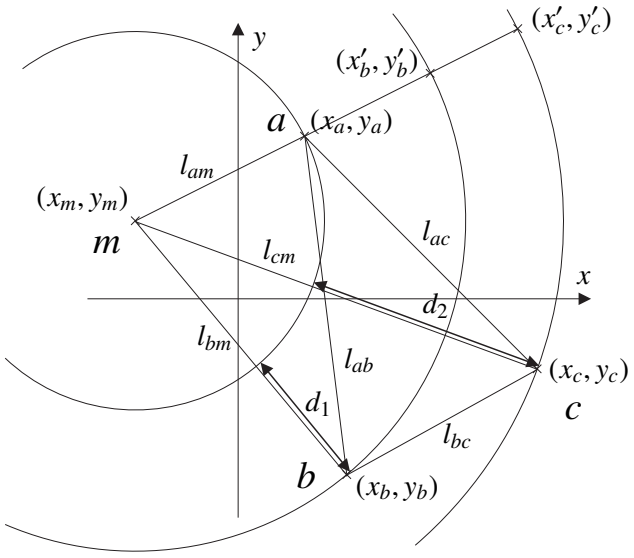


Figure B.1: Localization based on distance differences

The following dependencies can be set up:

$$\begin{aligned} d_1 &= l_{am} - l_{bm} \\ d_2 &= l_{am} - l_{cm} \end{aligned}$$

$$\begin{aligned}(x_m - x_a)^2 + (y_m - y_a)^2 &= {l_{am}}^2\\ (x_m - x_b)^2 + (y_m - y_b)^2 &= {l_{bm}}^2\\ (x_m - x_c)^2 + (y_m - y_c)^2 &= {l_{cm}}^2\end{aligned}$$

Note that two points at (x'_b, y'_b) and (x'_c, y'_c) have been introduced that yield the following equations:

$$\begin{aligned}(x_m - x'_b)^2 + (y_m - y'_b)^2 &= {l_{bm}}^2\\ (x_m - x'_c)^2 + (y_m - y'_c)^2 &= {l_{cm}}^2\end{aligned}$$

$$\begin{aligned}(x'_b - x_a)^2 + (y'_b - y_a)^2 &= {d_1}^2\\ (x'_c - x_a)^2 + (y'_c - y_a)^2 &= {d_2}^2\end{aligned}$$

$$\begin{aligned}x'_b &= \frac{-d_1(x_m - x_a)}{\sqrt{(x_m - x_a)^2 + (y_m - y_a)^2}} + x_a\\ y'_b &= \frac{-d_1(y_m - y_a)}{\sqrt{(x_m - x_a)^2 + (y_m - y_a)^2}} + y_a\\ x'_c &= \frac{-d_2(x_m - x_a)}{\sqrt{(x_m - x_a)^2 + (y_m - y_a)^2}} + x_a\\ y'_c &= \frac{-d_2(y_m - y_a)}{\sqrt{(x_m - x_a)^2 + (y_m - y_a)^2}} + y_a\end{aligned}$$

From these equations, we obtain

$$x_m = c_1 y_m + c_2 \tag{B.1}$$

with

$$\begin{aligned}c_1 &= \frac{-d_1\, y_a + d_2\, y_a - d_2\, y_b + d_1\, y_c}{c_3}\\ c_2 &= \frac{d_2\,\left(-{d_1}^2 - {x_a}^2 + {x_b}^2 - {y_a}^2 + {y_b}^2\right) + d_1\,\left({d_2}^2 + {x_a}^2 - {x_c}^2 + {y_a}^2 - {y_c}^2\right)}{2\, c_3}\\ c_3 &= d_1\, x_a - d_2\, x_a + d_2\, x_b - d_1\, x_c\end{aligned}$$

where x_a, y_a, x_b, y_b, x_c and y_c are known and d_1, d_2 given. From the same set of equations, a quadratic equation for y_m can be derived:

$$y_{m1,2} = -\frac{k_1}{2k_3} \pm \sqrt{\left(\frac{k_1}{2k_3}\right)^2 - \frac{k_2}{k_3}} \tag{B.2}$$

with

$$\begin{aligned}k_1 &= 4{d_1}^2\,(y_a + y_b) - 4\,(y_a - y_b)\,\left(-2\, c_2\, x_a + {x_a}^2 + 2\, c_2\, x_b - {x_b}^2 + {y_a}^2 - {y_b}^2\right)\\ &\quad -8c_1\, c_2\,\left({d_1}^2 - (x_a - x_b)^2\right) + 4c_1{d_1}^2\,(x_a + x_b) - 4c_1\,(x_a - x_b)\,\left({x_a}^2 - {x_b}^2 + {y_a}^2 - {y_b}^2\right)\\ k_2 &= {d_1}^4 - 4\,{c_2}^2\,\left({d_1}^2 - (x_a - x_b)^2\right) + \left({x_a}^2 - {x_b}^2 + {y_a}^2 - {y_b}^2\right)^2 - 2\,{d_1}^2\,\left({x_a}^2 + {x_b}^2 + {y_a}^2 + {y_b}^2\right)\\ &\quad +4\, c_2\, {d_1}^2\,(x_a + x_b) - 4c_2\,(x_a - x_b)\,\left({x_a}^2 - {x_b}^2 + {y_a}^2 - {y_b}^2\right)\\ k_3 &= 4\,\left(-{d_1}^2\left(1 + {c_1}^2\right) + (c_1\,(x_a - x_b) + y_a - y_b)^2\right)\end{aligned}$$

The result (x_m, y_m) can be obtained, apart from the square root in Eq. B.2, by employing standard linear algebra techniques.

Due to the quadratic relationship, three situations may occur: The position can be unambiguously (one real solution), ambiguously (two real solutions) or not be determined at all (one or two complex solutions). If two real solutions exist, two potential positions (x_{m1}, y_{m1}) and (x_{m2}, y_{m2}) are obtained. Nevertheless, in certain cases, the true position can be unambiguously found: This can be done by checking the distance differences d_1 and d_2 for the two positions using the following formulas, and subsequently comparing the results with the original (measured) values. These equations follow directly from the geometrical arrangement shown in Fig. B.1:

$$\begin{aligned} d_1(x_m, y_m) &= l_{am} - l_{bm} \\ &= \sqrt{(x_m - x_a)^2 + (y_m - y_a)^2} - \sqrt{(x_m - x_b)^2 + (y_m - y_b)^2} \qquad \text{(B.3)} \\ d_2(x_m, y_m) &= l_{am} - l_{cm} \\ &= \sqrt{(x_m - x_a)^2 + (y_m - y_a)^2} - \sqrt{(x_m - x_c)^2 + (y_m - y_c)^2} \qquad \text{(B.4)} \end{aligned}$$

List of Acronyms

AGPS ..... Assisted GPS

AGV ..... Automated guided vehicle

AOA ..... Angle of arrival

CIE ..... Commission Internationale de l'Eclairage

CMOS ..... Complementary metal-oxide semiconductor

COTS ..... Commercial off the shelf

DF ..... Direction finding

DFT ..... Discrete Fourier transform

DGPS ..... Differential GPS

DHT ..... Discrete Hartley transform

DOA ..... Direction of arrival

DSP ..... Digital signal processor

DTOA ..... Differential time of arrival

EMI ..... Electromagnetic interference

FFT ..... Fast Fourier transform

FOV ..... Field of view, field of vision

FPGA ..... Field programmable gate array

GNSS ..... Global navigation satellite system

IDFT ..... Inverse discrete Fourier transform

INS ..... Inertial navigation system

IR ..... Infrared

IrDA ..... Infrared Data Association

IRED ..... Infrared emitting diode

LAN ..... Local area network

LD ..... Laser diode

LED ..... Light emitting diode

LIDAR ..... Light detection and ranging

LPI ..... Low pass interpolation

LPS ..... Local positioning system

LSI ..... linear, shift invariant

NIR ..... Near IR

PDA ..... Personal digital assistant

RADAR ..... Radio detection and ranging

RFID ..... Radio frequency identification

RSS ..... Received signal strength

SGV ..... Self guided vehicle

SNR ..... Signal to noise ratio

SONAR ..... Sound navigation and ranging

TDOA ..... Time difference of arrival

TOA ..... Time of arrival

TOF ..... Time of flight

ULA ..... Uniform linear array

UWB ..... Ultra wideband

VFI ..... Virtual filter interpolation

VHDL ..... Very high speed integrated circuit hardware description language

Nomenclature

Symbols used in Chapter 4

Note that p and q denote place holders for one of the units a, b, c, d and m.

Symbols concerning the geometric arrangement

(x_a, y_a), (x_b, y_b),

(x_c, y_c), (x_d, y_d) ... Positions of reference units a, b, c and d

(x_m, y_m) ..... Position of the mobile target m

d_1, d_2 ..... Distance differences

l_{pq} ..... Physical distance between units p and q

Time measurement symbols (assuming ideal clocks)

t_g ..... Virtual global time

t_p ..... Local time of unit p

Δt_p ..... Clock offset between global time t_g and local time of unit p, t_p

t_{pq} ..... Physical signal propagation time between units p and q

$t_{s,p}$ ..... Local signal transmission time at unit p

$t_{r,pq}$ ..... Local signal reception time of a signal transmitted from unit p to unit q

v ..... Signal propagation speed

Symbols concerning the proposed measurement approach

T_p, T_q ..... Clock period at units p and q

σ_p ..... Clock jitter of the oscillator at unit p

$C_{q,pi}$ ..... Current counter value after the $(i-1)$-th pulse was detected

$t_{s0,p}$ ..... Delay until the emission of pulses from unit p is initiated

S'_{pq} ..... Sum of counter values $C_{q,pi}$

S_{pq} ..... Cumulative time value corresponding to the sum of counter values, S'_{pq}

N_a, N_b, N_c ..... Number of measurements where reference unit a, b or c act as transmitter, and the remaining two reference units as receivers

N_m ..... Number of measurements where the mobile unit m acts as transmitter, and the reference units as receivers

N_B ..... Number of burst measurements

T_B ..... Burst time

t_{SD} ..... Switch delay, i.e. the constant time span in which the measurement mode is switched. During this time span, no measurements can be performed.

T_{sim} ..... Overall simulation time

Symbols concerning the measurement performance

Δd_1, Δd_2 ..... Distance difference measurement errors

D_1, D_2 ..... *Measured* distance differences

Symbols used in Chapter 5

Symbols concerning the triangulation geometry

(x_a, y_a), (x_b, y_b) ... Positions of reference units a and b

(x_m, y_m) ..... Position of the mobile target m

α_a, α_b ..... Angle of arrival at reference units a and b

Sensor sensitivity symbols

κ ..... Parameter of the sensor sensitivity profile. If all sensors l have individual profiles, then instead of κ, the symbol κ_l is used.

σ_κ, σ_τ, σ_ρ, σ_a ... Measure of the individual irregularities of the sensor sensitivity profiles

τ ..... Parameter of the sensor sensitivity profile. If all sensors l have individual profiles, then instead of τ, the symbol τ_l is used.

Symbols specifying the detector array

γ_{FOV} ..... Half angle of the field of view of the detector array

Φ ..... Constant angular displacement between two sensors. Φ defines the sampling period used in the low pass interpolation approach. Φ is only defined if two adjacent sensors are rotated against each other by equal amounts.

ρ_l ..... Angular displacement of sensor l from the optical centerline of the detector array

$\boldsymbol{\rho}$ ..... Vector of length L containing all ρ_l

L ..... Number of single sensors in the detector array

Signals in the continuous domain

$d(\phi)$ ..... General sensitivity profile of a single sensor

$d_l(\phi)$ ..... Shifted sensitivity profile of sensor l, according to the sensor's angular displacement ρ_l

$b(\phi)$ ..... Radiant intensity profile incident on the detector array

$b'(\phi)$ ..... Reconstructed radiation profile, after passing the low pass filter

$b''(\phi)$ ..... Reconstructed radiation profile, after passing the low pass filter and the sensitivity profile compensation filter

$x(\phi)$ ..... Continuous representation of the sensor readings at angles ϕ. $x(\phi)$ is non-zero only at the angular positions of the L sensors, ρ_l.

$H_{LP}(f)$ ..... Transfer function of the reconstruction low pass filter

Symbols and signals in the discrete domain

$\mathbf{d}$, $d(n)$ ..... Sensitivity profile of a single sensor

$\mathbf{d}_l$, $d_l(n)$ ..... Sensitivity profile of sensor l. $d_l(n)$ complies with $d(n)$, except that the profile is shifted to the angular position of sensor l.

$\underline{\mathbf{D}}$ ..... Sensitivity profile matrix, containing all shifted sensor sensitivities $\mathbf{d}_l$

$\mathbf{b}$, $b(n)$ ..... Radiation profile incident on the detector array

$b'(n)$ ..... Reconstructed radiation profile, after the low pass interpolation filter

$b''(n)$ ..... Reconstructed radiation profile, after the low pass interpolation filter and the sensitivity profile compensation filter

$\mathbf{h_{LP}}$, $h_{LP}(n)$ ..... Impulse response of the low pass interpolation filter

$\mathbf{H_{LP}}$, $H_{LP}(k)$ ..... Transfer function of the low pass interpolation filter

x_l ..... Reading of sensor l

$\mathbf{x}$ ..... Vector containing all sensor readings x_l

$\mathbf{h}_0$, $h_0(n)$ ..... Impulse response of the *nominal* virtual filter

$\underline{\mathbf{H}}_0$, $\mathbf{H}_0$, $H_0(k)$ .... Transfer matrix/function of the *nominal* virtual filter

$\underline{\mathbf{H}}$ ..... Transfer matrix of the *approximated* virtual filter

$\mathbf{v}_0$, $v_0(n)$ ..... Output signal of the *nominal* virtual filter

$\mathbf{v}$, $v(n)$ ..... Output signal of the *approximated* virtual filter

$\underline{\mathbf{\Omega}}$ ..... Weighting matrix consisting of the weighting coefficients $\omega_n(l)$

$\mathbf{s}$, $s(n)$ ..... Reconstructed radiation profile using the virtual filter interpolation method

Δr ..... Angular resolution of the discrete approach

N ..... Sequence length corresponding to the entire angular range

F ..... Sequence length corresponding to the half angle of the field of view, γ_{FOV}

Symbols concerning the measurement performance

α ..... *True* angle of arrival

$\alpha_{b'}$ ..... *Measured* angle of arrival using the low pass interpolation method

α_s ..... *Measured* angle of arrival using the virtual filter interpolation method

$\epsilon_{b'}(\alpha)$ ..... Absolute AOA error of the measured angle $\alpha_{b'}$ compared to the true angle α

$\epsilon_s(\alpha)$ ..... Absolute AOA error of the measured angle α_s compared to the true angle α

G ..... Number of sample sets $\{x_l(g)\}$, $0 \leq g < G$

Complexity symbols

A_N ..... Number of additions at a given sequence length N

M_N ..... Number of multiplications at a given sequence length N

Bibliography

[1] *IEEE Std 802.11-1999 Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications*.

[2] M. Addlesee, A. Jones, F. Livesey, and F. Samaria. The ORL active floor. *IEEE Personal Communications*, 4(5):35–41, October 1997.

[3] American National Standards Institute. *ANSI-Z136-1: American National Standard for the Safe Use of Lasers*, 2000.

[4] P. Bahl and V. N. Padmanabha. RADAR: An in-building RF-based user location and tracking system. In *19th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, volume 2, pages 775–784, Tel Aviv, Israel, March 2000.

[5] Y. Bar-Shalom, L. X. R., and T. Kirubarajan. *Estimation with Applications to Tracking and Navigation*. John Wiley, 2001.

[6] J. Barnes, C. Rizos, M. Kanli, D. Small, G. Voight, N. Gambale, J. Lamance, T. Nunan, and C. Reid. Indoor industrial machine guidance using Locata: A pilot study at BlueScope Steel. In *60th Annual Meeting of the U.S. Institute of Navigation*, pages 533–540, Dayton, OH, USA, June 2004.

[7] J. Barnes, C. Rizos, J. Wang, D. Small, G. Voight, and N. Gambale. High precision indoor and outdoor positioning using LocataNet. In *International Symposium on GPS/GNSS*, pages 9–18, Tokyo, Japan, November 2003.

[8] J. Barnes, J. Wang, C. Rizos, D. Small, G. Voight, and N. Gambale. Locatanet: Intelligent synchronised pseudolites for cm-level stand-alone positioning. In *11th International Association of Institutes of Navigation (IAIN) World Congress, CD-ROM*, Berlin, Germany, October 2003.

[9] J. R. Barry. *Wireless infrared communications*. Kluwer, Boston, MA, USA, 1994.

[10] J. R. Barry, J. M. Kahn, W. J. Krause, E. A. Lee, and D. G. Messerschmitt. Simulation of multipath impulse response for indoor wireless optical channels. *IEEE Journal on Selected Areas in Communications*, 11(3):367–379, April 1993.

[11] J. R. Barry, E. A. Lee, and D. G. Messerschmitt. *Digital Communication*. Kluwer Academic Publishers, Boston, 3rd edition, 2004.

[12] X. Bian, G. D. Abowd, and J. M. Rehg. Using sound source localization in a home environment. In *3rd International Conference on Pervasive Computing (Pervasive)*, pages 19–36, Munich, Germany, May 2005.

[13] J. Borenstein, H. R. Everett, and L. Feng. Where am I? Sensors and methods for mobile robot positioning. Technical report, University of Michigan, MI, USA, April 1996.

[14] J. Borenstein and L. Feng. Gyrodometry: A new method for combining data from gyros and odometry in mobile robots. In *IEEE International Conference on Robotics and Automation*, pages 423–428, Minneapolis, MN, USA, April 1996.

[15] A. C. Boucouvalas. Ambient light noise and its effect on receiver design for indoor optical wireless links. In *IEEE International Conference on Communications (ICC)*, volume 3, pages 1472–1476, Dallas, TX, USA, June 1996.

[16] I. N. Bronshtein, K. A. Smendyayev, G. Musiol, and H. Muehlig. *Handbook of Mathematics*. Springer, Berlin, 2004.

[17] C. G. Broyden. The convergence of a class of double-rank minimization algorithms. *Journal of the Institute of Mathematics and its Applications*, 6:76–90, 1970.

[18] J. J. Caffery and G. L. Stüber. Overview of radiolocation in CDMA cellular systems. *IEEE Communications Magazine*, 36(4):38–45, April 1998.

[19] J. B. Carruthers. *Encyclopedia of Telecommunications*, chapter Wireless Infrared Communications. Wiley, New York, USA, 2002.

[20] J. B. Carruthers and J. M. Kahn. Modeling of nondirected wireless infrared channels. *IEEE Transactions on Communications*, 45(10):1260–1268, 1997.

[21] Y. T. Chan and K. C. Ho. A simple and efficient estimator for hyperbolic location. *IEEE Transactions on Signal Processing*, 42(8):1905–1915, August 1994.

[22] Commission Internationale de l'Eclairage (CIE) and International Electrotechnical Commission (IEC), Austria, Vienna. *International Lighting Vocabulary*, 4th edition, 1987.

[23] I. D. Coope. Reliable computation of the points of intersection of n spheres in $\mathbb{R}^n$. *Australian & New Zealand Industrial and Applied Mathematics Journal*, 42 (E):C461–C477, 2000.

[24] C. Drane, M. Macnaughtan, and C. Scott. Positioning GSM telephones. *IEEE Communications Magazine*, 36(4):46–54, 59, April 1998.

[25] P. Duhamel and M. Vetterli. Cyclic convolution of real sequences: Hartley versus fourier and new schemes. In *IEEE Conference on Acoustics, Speech and Signal Processing*, volume 11, pages 229–232, Tokyo, Japan, April 1986.

[26] E. Elnahrawy, X. Li, and R. P. Martin. The limits of localization using signal strength: A comparative study. In *1st IEEE International Conference on Sensor and Ad hoc Communications and Networks (SECON)*, Santa Clara, CA, USA, October 2004.

[27] B. T. Fang. Simple solutions for hyperbolic and related position fixes. *IEEE Transactions on Aerospace and Electronic Systems*, 26(5):748–753, 1990.

[28] S. Feldmann, K. Kyamakya, A. Zapater, and Z. Lue. An indoor Bluetooth-based positioning system: concept, implementation and experimental evaluation. In *International Conference on Wireless Networks (ICWN)*, pages 109–113, Las Vegas, NV, USA, June 2003.

[29] J. Fletcher and M. J. D. Powell. A rapidly convergent descent method for minimzation. *Computer Journal*, 6:163–168, 1963.

[30] R. J. Fontana. Experimental results from an ultra wideband precision geolocation system. In *Ultra-Wideband, Short-Pulse Electromagnetics 5*, pages 215–224. Kluwer, NY, 2000.

[31] R. J. Fontana and S. J. Gunderson. Ultra-wideband precision asset location system. In *IEEE Conference on Ultra Wideband Systems and Technologies (UWBST)*, pages 147–150, Baltimore, MD, USA, May 2002.

[32] R. J. Fontana, E. Richley, and J. Barney. Commercialization of an ultra wideband precision asset location system. In *IEEE Conference on Ultra Wideband Systems and Technologies*, pages 369–373, Reston, VA, USA, November 2003.

[33] D. Fox, J. Hightower, L. Liao, D. Schulz, and G. Borriello. Bayesian filtering for location estimation. *IEEE Pervasive Computing*, 2(3):24–33, July 2003.

[34] W. Foy. Position-location solutions by taylor-series estimation. *IEEE Transactions on Aerospace and Electronic Systems*, 12(2):187–193, 1976.

[35] J. Fraden. *Handbook of modern sensors – physics, design and applications*. Springer, 3rd edition, 2004.

[36] B. Friedlander. A passive localization algorithm and its accurancy analysis. *IEEE Journal of Oceanic Engineering*, 12(1):234–245, 1987.

[37] L. Girod and D. Estrin. Robust range estimation using acoustic and multimodal sensing. In *IEEE Intelligent Robots and Systems (IROS)*, volume 3, pages 1312–1320, Maui, HI, USA, November 2001.

[38] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, 1996.

[39] D. Hallaway, T. Hollerer, and S. Feiner. Coarse, inexpensive, infrared tracking for wearable computing. In *7th IEEE International Symposium on Wearable Computers (ISWC)*, pages 69–78, White Plains, NY, USA, October 2003.

[40] J. Hallberg, M. Nilsson, and K. Synnes. Positioning with Bluetooth. In *10th International Conference on Telecommunications (ICT)*, volume 2, pages 954–958, Tahiti, French Polynesia, March 2003.

[41] Hewlett-Packard. Fundamentals of Quartz Oscillators, Application Note 200-2, 1997.

[42] J. Hightower and G. Borriello. Location systems for ubiquitous computing. *IEEE Computer*, 34(8):57–66, August 2001.

[43] J. Hightower, R. Want, and G. Borriello. SpotON: An indoor 3d location sensing technology based on RF signal strength. UW CSE 00-02-02, University of Washington, Department of Computer Science and Engineering, Seattle, WA, USA, February 2000.

[44] HL Planar. *Specification of the Thermopile IR Sensor Unit TSE 01/08 L, Rev. 1.0.*

[45] B. Hofmann-Wellenhof, H. Lichtenegger, and J. Collins. *Global Positioning System: Theory and Practice*. Springer, 4th edition, 1997.

[46] R. D. Hudson. *Infrared System Engingeering*. Wiley series in pure and applied optics. Wiley, New York, USA, 1969.

[47] International Electrotechnical Commission (IEC). *EN60825-1 Ed. 1.2: Safety of laser products - Part 1: Equipment classification, requirements and user's guide*, 2001.

[48] B. Iske, B. Jäger, and U. Rückert. A ray-tracing approach for simulating recognition abilitites of active infrared sensor arrays. *IEEE Sensors Journal*, 4(2):237–247, April 2004.

[49] A. R. Jiménez. Absolute positioning systems. Technical Report SAM #003/1999, Instituto de Automática Industrial - CSIC (Council for Scientific Research), Madrid, Spain, May 1999.

[50] A. R. Jiménez, F. Seco, R. Ceres, and L. Calderón. Absolute localization using active beacons: A survey and IAI-CSIC contributions. White Paper, Instituto de Automática Industrial - CSIC (Council for Scientific Research), Madrid, Spain, 2004.

[51] Y. Kaddoura, J. King, and A. Helal. Cost-precision tradeoffs in unencumbered floor-based indoor location tracking. In *3rd International Conference On Smart homes and health Telematic (ICOST)*, Mogag, Canada, July 2005.

[52] J. M. Kahn and J. R. Barry. Wireless infrared communications. *Proceedings of the IEEE*, 85(2):265–298, February 1997.

[53] J. M. Kahn, W. J. Krause, and J. B. Carruthers. Experimental characterization of non-directed indoor infrared channels. *IEEE Transactions on Communications*, 43(2/3/4):1613–1623, 1995.

[54] J. M. Kahn, R. You, P. Djahani, A. G. Weisbin, B. K. Teik, and A. Tang. Imaging diversity receivers for high-speed infrared wireless communication page(s): 88-94. *IEEE Communications Magazine*, 36(12):88–94, December 1998.

[55] Fibre Channel – Methodologies for Jitter Specification, T11.2/Project 1230/ Rev 10, June 9, 1999.

[56] K. J. Krizman, T. E. Biedka, and T. S. Rappaport. Wireless position location: Fundamentals, implementation strategies, and sources of error. In *IEEE 47th Vehicular Technology Conference*, volume 2, pages 919–923, Phoenix, AZ, USA, May 1997.

[57] A. Krohn, M. Beigl, M. Hazas, H. Gellersen, and A. Schmidt. Using fine-grained infrared positioning to support the surface-based activities of mobile users. In *5th International Workshop on Smart Appliances and Wearable Computing (IWSAWC)*, Columbus, OH, USA, June 2005.

[58] G. Lachapelle. GNSS indoor location technologies. *Journal of Global Positioning Systems*, 3(1–2):2–11, 2004.

[59] B. Li, A. G. Dempster, C. Rizos, and J. Barnes. Hybrid method for localization using WLAN. In *Spatial Sciences Conference*, pages –, Melbourne, Australia, September 2005.

[60] D. E. Manolakis. Efficient solution and performance analysis of 3-d position estimation by trilateration. *IEEE Transactions on Aerospace and Electronic Systems*, 32(4):1239–1248, 1996.

[61] W. R. McCluney. *Introduction to Radiometry and Photometry.* Artech House, 1994.

[62] G. Mellen II., M. Pachter, and J. Raquet. Closed-form solution for determining emitter location using time difference of arrival measurements. *IEEE Transactions on Aerospace and Electronic Systems*, 39(3):1056–1058, July 2003.

[63] M. Minami, K. Hirasawa, H. Morikawa, and T. Aoyama. Implementation and evaluation of a distributed ultrasonic positioning system. In *1st International Workshop on Networked Sensing Systems (INSS)*, pages 75–78, Tokyo, Japan, June 2004.

[64] A. J. C. Moreira, R. T. Valadas, and A. M. de Oliveira Duarte. Characterisation and modelling of artificial light interference in optical wireless communication systems. In *6th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, volume 1, pages 326–331, Toronto, Canada, September 1995.

[65] A. J. C. Moreira, R. T. Valadas, and A. M. de Oliveira Duarte. Optical interference produced by artificial light. *Wireless Networks*, 3(2):131–140, 1997.

[66] K. Muthukrishnan, M. E. Lijding, and P. J. M. Havinga. Towards smart surroundings: Enabling techniques and technologies for localization. In *1st International Workshop on Location- and Context-Awareness (LoCA)*, volume LNCS 3479, pages 350–362, Oberpfaffenhofen, Germany, May 2005. Springer, Berlin.

[67] L. M. Ni, Y. Liu, Y. C. Lau, and A. P. Patil. LANDMARC: Indoor location sensing using active RFID. In *First IEEE International Conference on Pervasive Computing and Communications*, pages 407–415, Fort Worth, TX, USA, March 2003.

[68] A. V. Oppenheim and R. W. Schafer. *Discrete-Time Signal Processing*. Prentice Hall, London, UK, 2nd edition, 1999.

[69] I. Oppermann, M. Hämäläinen, and J. Iinatti. *UWB – Theory and Applications*. Wiley, 2004.

[70] R. J. Orr and G. D. Abowd. The smart floor: a mechanism for natural user identification and tracking. In *Conference on Human Factors in Computing Systems archive (CHI)*, pages 275–276, The Hague, The Netherlands, April 2000.

[71] N. B. Priyantha, A. Chakraborty, and H. Balakrishnan. The Cricket location-support system. In *6th International Conference on Mobile Computing and Networking (MobiCom)*, pages 32–43, Boston, MA, USA, August 2000.

[72] N. B. Priyantha, A. K. L. Miu, H. Balakrishnan, and S. Teller. The cricket compass for context-aware mobile applications. In *7th Annual International Conference on Mobile Computing and Networking (MOBICOM)*, pages 1–14, Rome, Italy, July 2001.

[73] C. Randell and H. L. Muller. Low cost indoor positioning system. In *3rd International Conference on Ubiquitous Computing (UbiComp)*, pages 42–48, Atlanta, GA, USA, October 2001.

[74] T. Rappaport. *Wireless Communications Ű Principles & Practice*. Prentice Halle, 2nd edition, 2001.

[75] A. Savvides, C.-C. Han, and M. B. Strivastava. Dynamic fine-grained localization in ad-hoc networks of sensors. In *ACM SIGMOBILE 7th Annual International Conference on Mobile Computing and Networking (MobiCom)*, pages 166–179, Rome, Italy, July 2001.

[76] H. Schau and A. Robinson. Passive source localization employing intersecting spherical surfaces from time-of-arrival differences. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 35(8):1223–1225, August 1987.

[77] R. Schmidt. Multiple emitter location and signal parameter estimation. *IEEE Transactions on Antennas and Propagation*, AP-34(3):276–280, 1986.

[78] A. Shitzer and R. Eberhart, editors. *Heat Transfer in Medicine and Biology: Analysis and Applications*, volume 2. Plenum Publishing Corporation, London, UK, 1985.

[79] C. Singh, J. John, Y. N. Singh, and K. K. Tripathi. A review of optical wireless systems. *IETE Technical Review*, 19(1):3–17, Jan.–Feb. 2002.

[80] A. Smith, H. Balakrishnan, M. Goraczko, and N. Priyantha. Tracking moving devices with the cricket location system. In *2nd international conference on Mobile systems, applications, and services (MobiSys)*, pages 190–202, Boston, MA, USA, June 2004.

[81] J. O. Smith and J. S. Abel. Closed-form least-squares source location estimation from range-difference measurements. *IEEE Transactions on Acoustics, Speech and Signal Processing*, ASSP-35(12):1661–1669, 1987.

[82] H. Sorensen, D. Jones, M. Heideman, and C. Burrus. Real-valued fast fourier transform algorithms. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 35(6):849–863, 1987.

[83] A. M. Street, P. N. Stavrinou, D. C. O'Brien, and D. J. Edwards. Indoor optical wireless systems - a review. *Optical and Quantum Electronics*, 29(3):349–378, January 1997.

[84] A. P. Tang, J. M. Kahn, and K.-P. Ho. Wireless infrared communication links using multi-beam transmitters and imaging receivers. In *IEEE International Conference on Communications (ICC)*, volume 1, pages 180–186, 1996.

[85] J. A. Tauber. Indoor location systems for pervasive computing. Technical report, Massachusetts Institute of Technology, Cambridge, MA, USA, 2002.

[86] D. J. Torrieri. Statistical theory of passive location systems. *IEEE Transactions on Aerospace and Electronic Systems*, 20(2):183–198, 1984.

[87] Ubisense. Smart Space. Brochure, http://www.ubisense.net/Files/PDF/Ubisense_brochure.pdf (confirmed August 2005), 2004.

[88] F. van Diggelen and C. Abraham. Indoor GPS technology. In *CTIA Wireless Agenda*, Dallas, TX, USA, May 2001.

[89] A. W. van Herwaarden, D. C. van Duyn, B. W. van Oudheusden, and P. M. Sarro. Integrated thermopile sensors. *Sensors and Actuators*, A21–A23:621–630, 1989.

[90] Vishay Telefunken. *About IrDA. IrDA-Compatible Data Transmission. IrDa Standard – Physical Layer*, 05.01. Document Number 82500.

[91] H.-G. Wagemann. *Grundlagen der optoelektronischen Halbleiterbauelemente*. Teubner, Stuttgart, Germany, 1998.

[92] R. Want, A. Hopper, V. Falcão, and J. Gibbons. The active badge location system. *ACM Transactions on Office Information Systems (TOIS)*, 10(1):91–102, January 1992.

[93] A. Ward, A. Jones, and A. Hopper. A new location technique for the Active Office. *IEEE Personal Communications*, 4(5):42–47, October 1997.

[94] J. Werb and C. Lanzl. Designing a positioning system for finding things and people indoors. *IEEE Spectrum*, 35(9):71–78, 1998.

[95] Z. Xiang, S. Song, J. Chen, H. Wang, and J. Huang. A wireless LAN-based indoor positioning technology. *IBM Journal of Research and Development*, 48(5/6):617–626, September/October 2004.

[96] G. Yun and M. Kavehrad. Spot-diffusing and fly-eye receivers for indoor infrared radio communications. In *IEEE International Conference on Selected Topics in Wireless Communications*, pages 262–265, Vancouver, BC, Canada, June 1992.