Computer Vision without Vision

Methods and Applications of Radio and Audio Based SLAM

Batstone, Kenneth John

2020

Document Version:
Publisher's PDF, also known as Version of record

– CENTRUM SCIENTIARUM MATHEMATICARUM –

Computer Vision without Vision

Methods and Applications of Radio and Audio Based SLAM

KENNETH JOHN BATSTONE

Lund University
Faculty of Engineering
Centre for Mathematical Sciences
Mathematics

Printed by Media-Tryck, Lund 2020 NORDIC SWAN ECOLABEL 3041 0903

Doctoral Thesis in Mathematical Sciences 2020:3
ISBN 978-91-7895-620-3
LUTFMA-1069-2020
ISSN 1404-0034

Computer Vision without Vision

Computer Vision without Vision

Methods and Applications of Radio and Audio Based SLAM

by Kenneth John Batstone

LUND
UNIVERSITY

Thesis for the degree of Doctor of Philosophy
Thesis advisors: Prof. Kalle Åström, Dr. Magnus Oskarsson,
Prof. Fredrik Tufvesson, Prof. Bo Bernhardsson
Faculty opponent: Prof. Heidi Kuusniemi, University of Vaasa, Finland

To be presented, with the permission of the Faculty of Engineering of Lund University, for public criticism in lecture hall MH:Gårding at the Centre for Mathematical Sciences on Friday, the 2nd of October 2020 at 13:15.

Organization	Document name
LUND UNIVERSITY Centre for Mathematical Sciences Box 118 SE–221 00 LUND Sweden	DOCTORATE THESIS IN MATHEMATICAL SCIENCES Date of disputation 2020-10-02
Author(s) Kenneth John Batstone	Sponsoring organization

Title and subtitle

Computer Vision without Vision: Methods and Applications of Radio and Audio Based SLAM

Abstract

The central problem of this thesis is estimating receiver-sender node positions from measured receiver-sender distances or equivalent measurements. This problem arises in many applications such as microphone array calibration, radio antenna array calibration, mapping and positioning using ultra-wideband and mapping and positioning using round-trip-time measurements between mobile phones and Wi-Fi-units. Previous research has explored some of these problems, creating minimal solvers for instance, but these solutions lack real world implementation. Due to the nature of using different media, finding reliable receiver-sender distances is tough, with many of the measurements being erroneous or to a worse extent missing. Therefore in this thesis, we explore using minimal solvers to create robust solutions, that encompass small erroneous measurements and work around missing and grossly erroneous measurements.

This thesis focuses mainly on Time-of-Arrival measurements using radio technologies such as Two-way-Ranging in Ultra-Wideband and a new IEEE standard 802.11mc found on many WiFi modules. The methods investigated, also related to Computer Vision problems such as Stucture-from-Motion. As part of this thesis, a range of new commercial radio technologies are characterised in terms of ranging in real world enviroments. In doing so, we have shown how these technologies can be used as a more accurate alternative to the Global Positioning System in indoor enviroments. Further to these solutions, more methods are proposed for large scale problems when multiple users will collect the data, commonly known as Big Data. For these cases, more data is not always better, so a method is proposed to try find the relevant data to calibrate large systems.

Key words

TOA, Self-Calibration, Localization, 802.11mc, Round-Trip Time

Classification system and/or index terms (if any)

Supplementary bibliographical information		Language
		English
ISSN and key title 1404-0034		ISBN 978-91-7895-620-3 (print) 978-91-7895-621-0 (pdf)
Recipient's notes	Number of pages 190	Price
	Security classification	

Computer Vision without Vision

Methods and Applications of Radio and Audio Based SLAM

by Kenneth John Batstone

LUND
UNIVERSITY

A doctoral thesis at a university in Sweden takes either the form of a single, cohesive research study (monograph) or a summary of research papers (compilation thesis), which the doctoral student has written alone or together with one or several other author(s).

In the latter case the thesis consists of two parts. An introductory text puts the research work into context and summarizes the main points of the papers. Then, the research publications themselves are reproduced, together with a description of the individual contributions of the authors. The research papers may either have been already published or are manuscripts at various stages (in press, submitted, or in draft).

Cover illustration front: A Picture of the code used in this thesis. (Credits: Kenneth Batstone).

Funding information: The thesis work was financially supported by Mobile and Pervasive Computing Institute (MAPCI), Excellence Center at Linköping - Lund in Information Technology (ELLIIT) and the Royal Physiographic Society in Lund (Kungliga Fysiografiska Sällskapet i Lund)

Faculty of Engineering, Centre for Mathematical Sciences
Doctoral Thesis in Mathematical Sciences 2020:3
LUTFMA-1069-2020
ISBN: 978-91-7895-620-3 (print)
ISBN: 978-91-7895-621-0 (pdf)
ISSN: 1404-0034

Printed in Sweden by Media-Tryck, Lund University, Lund 2020

There is nothing noble in being superior to your fellow man; true nobility is being superior to your former self.

Ernest Hemingway

Contents

Abstract

The central problem of this thesis is estimating receiver-sender node positions from measured receiver-sender distances or equivalent measurements. This problem arises in many applications such as microphone array calibration, radio antenna array calibration, mapping and positioning using ultra-wideband and mapping and positioning using round-trip-time measurements between mobile phones and Wi-Fi-units. Previous research has explored some of these problems, creating minimal solvers for instance, but these solutions lack real world implementation. Due to the nature of using different media, finding reliable receiver-sender distances is tough, with many of the measurements being erroneous or to a worse extent missing. Therefore in this thesis, we explore using minimal solvers to create robust solutions, that encompass small erroneous measurements and work around missing and grossly erroneous measurements.

This thesis focuses mainly on Time-of-Arrival measurements using radio technologies such as Two-way-Ranging in Ultra-Wideband and a new IEEE standard 802.11mc found on many WiFi modules. The methods investigated, also related to Computer Vision problems such as Stucture-from-Motion. As part of this thesis, a range of new commercial radio technologies are characterised in terms of ranging in real world enviroments. In doing so, we have shown how these technologies can be used as a more accurate alternative to the Global Positioning System in indoor enviroments. Further to these solutions, more methods are proposed for large scale problems when multiple users will collect the data, commonly known as Big Data. For these cases, more data is not always better, so a method is proposed to try find the relevant data to calibrate large systems.

Popular Summary

Society has always been dependent on good navigation. From using the stars to navigate the seas, or simply just remembering landmarks to help you go back to an area where there is a plentiful source of food. These systems of navigation still exist today but with more urbanised areas and a vastly greater number of people, the demand and precision has also increased. In modern society, the vast majority of mobile phones, cars and planes all have a high requirement of good navigation and positioning. For mobile phone users, it can help you find your way to a specific shop, find local services you may require and also give your location to others in an emergency. For cars, it can help you find your way from one place to another and update the route depending on traffic, or to avoid tolls for instance.

One of the main systems that is used today, is Global Positioning System (GPS). This system has been effective on meeting the high demands of the users, but it comes with its drawbacks. In urban areas, where the majority of people today live, the buildings can block the signals from the GPS Satellites. This in turn means the positioning of users can become inaccurate.

So how can we overcome this issue? This is the key question of this thesis. Here, we explore using other radio based systems to help navigate and position indoors. One key infrastructure which already exists in urban areas is Wi-Fi. Most homes and offices, even telephone boxes, have a Wi-Fi router in them. As part of our work, we have developed methods to find these routers in order to use them as landmarks to help navigate indoors. One of the main issues we have incurred due to the complex indoor environment, is that Wi-Fi measurements we use can also become inaccurate. In our methods we try and find a model in the data in order to identify the bad measurements. By doing so we can ignore the bad measurements and accurately find the locations of such Wi-Fi routers.

We also looked at new radio technologies to further increase our precision. These advancements have the potential to help society greatly, and meet the increased needs of navigation in the future. As we come into the new era of wireless devices, commonly known as 5G, more and more devices will require positioning and navigation. This generation of wireless devices has been developed to meet the demands of the Internet of things (IoT). IoT is

the idea of objects with internet capabilities, such as a toaster. These IoT devices usually have some form of radio based communication system, such as Wi-Fi, Bluetooth or Ultra-Wideband (UWB). Although a toaster may not have a need for positioning, IoT devices could be installed on robots, hospital equipment and even packages. Having the ability to find or help these objects to navigate is important and advantageous to society. In this thesis, we also explore using our previous methods to find these radio based devices, in an environment where there are many devices and multiple users to help give insight to future requirements.

Populärvetenskaplig Sammanfattning

Samhället har alltid varit beroende av fungerande navigering. Från att använda stjärnorna för att navigera över haven, eller helt enkelt bara komma ihåg landmärken för att hjälpa dig återvända till ett område där det finns rikligt med mat. Dessa navigationssystem finns fortfarande idag men med mer urbaniserade områden och ett mycket större antal människor har efterfrågan och precisionen också ökat. I det moderna samhället har de allra flesta mobiltelefoner, bilar och flygplan ett stort krav på väl fungerande navigering och positionering. För mobiltelefonanvändare kan det hjälpa dig att hitta vägen till en specifik butik, hitta lokala tjänster som du kan behöva och även ge din plats till andra i en nödsituation. För bilar kan det hjälpa dig att hitta från en plats till en annan och uppdatera rutten beroende på trafik eller att undvika vägtullar.

Ett av de viktigaste systemen som används idag är GPS (GPS). Detta system har varit effektivt för att möta användarnas höga krav, men det har sina nackdelar. I stadsområden, där majoriteten av människor bor idag, kan byggnaderna blockera signalerna från GPS-satelliterna. Detta innebär i sin tur att användarnas placering kan bli felaktig.

Så hur kan vi bemöta det här problemet? Detta är den viktigaste frågan i den här avhandlingen. Här utforskar vi andra radiobaserade system för att hjälpa till att navigera inomhus. En viktig infrastruktur som redan finns i stadsområden är Wi-Fi. De flesta hem och kontor, även telefonlådor, har en Wi-Fi-router i sig. Som en del av vårt arbete har vi utvecklat metoder för att hitta dessa routrar för att använda dem som landmärken för att navigera inomhus. Ett av de största problemen vi har haft på grund av den komplexa inomhusmiljön är att Wi-Fi-mätningar vi använder också kan bli felaktiga. I våra metoder försöker vi hitta en modell i uppgifterna för att identifiera de dåliga mätningarna. Genom att göra det kan vi ignorera de dåliga mätningarna och hitta exakt platserna för sådana Wi-Fi-routrar.

Vi tittade också på nya radiotekniker för att ytterligare öka vår precision. Dessa framsteg har potential att hjälpa samhället kraftigt och möta det ökade navigationsbehovet i framtiden. När vi kommer in i den nya eran med trådlösa enheter, allmänt känd som 5G, kommer fler och fler enheter att kräva positionering och navigering. Denna generation trådlösa enheter har utvecklats för att uppfylla kraven från Internet of Things (IoT). IoT är idén att ha objekt

med internetfunktioner, till exempel en brödrost. Dessa IoT-enheter har vanligtvis någon form av radiobaserat kommunikationssystem, till exempel Wi-Fi, Bluetooth eller Ultra-Wideband (UWB). Även om en brödrost kanske inte har något behov av positionering, kan IoT-enheter installeras på robotar, sjukhusutrustning eller till och med i paket. Att ha förmågan att hitta eller hjälpa dessa objekt att navigera är viktigt och fördelaktigt för samhället. I denna avhandling undersöker vi också våra tidigare metoder för att hitta dessa radiobaserade enheter, i en miljö där det finns många enheter och flera användare som kan ge insikt om framtida krav på tekniken.

Acknowledgements

I would like to thank my supervisors, Prof. Kalle Åström, Dr. Magnus Oskarsson, Prof. Fredrik Tufvesson and Prof. Bo Bernhardsson, for giving me the opportunity to study for a Ph.D. It has been a life long ambition of mine to obtain a Ph.D. so the opportunity is greatly appreciated. In particular I would like to thank Kalle and Magnus, for always being available for questions and willing to listen to my ideas, regardless if they are good or not. Also I would like to thank my colleagues at the Centre of Mathematical Sciences, in particular the Computer Vision Group for all their support. A big thank you goes to Björn Lindquist and the rest of the team at Combain AB for giving me much of the resources for the experiments and a great insight into real world indoor navigation problems.

I want to thank my friends for countless hours of fun and down time. I would also like to thank my family, Batstones and Dorozynskis, for all the support over the years.

Finally, I wish to thank my partner, Karolina Dorozynska, for pointing me in the right direction when I needed it and supporting what I do tirelessly.

List of Publications

This thesis is based on the following publications, referred to by their Roman numerals:

I **Robust Time-of-Arrival Self Calibration and Indoor Localization using Wi-Fi Round-Trip Time Measurements**
K. Batstone, M. Oskarsson, K. Åström
2016 IEEE International Conference on Communications Workshops (ICC), Kuala Lumpur, 2016, pp. 26-31.

II **Robust Time-of-Arrival Self Calibration with Missing Data and Outliers**
K. Batstone, M. Oskarsson, K. Åström
2016 24th European Signal Processing Conference (EUSIPCO), Budapest, 2016, pp. 2370-2374.

III **Robust Self-Calibration of Constant Offset Time-Difference-of-Arrival**
K. Batstone, G. Flood, T. Beleyur, V. Larsson, H. R. Goerlitz, M. Oskarsson, K. Åström
2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, United Kingdom, 2019, pp. 4410-4414.

IV **Trust No One: Low Rank Matrix Factorization Using Hierarchical RANSAC**
M. Oskarsson, K. Batstone, K. Åström
IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 5820-5829.

V **Towards Real-time Time-of-Arrival Self-Calibration using Ultra-Wideband Anchors**
K. Batstone, M. Oskarsson, K. Åström
2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN), Sapporo, 2017, pp. 1-8.

VI **Collaborative Merging of Radio SLAM Maps in View of Crowd-sourced Data Acquisition and Big Data**
K. Batstone, M. Oskarsson, K. Åström
In Proceedings of the 8th International Conference on Pattern Recognition Applications and Methods (2019)- Volume 1: ICPRAM, ISBN 978-989-758-351-3, pages 807-813.

VII **Robust Phase-Based Positioning Using Massive MIMO With Limited Bandwidth**
X. Li , K. Batstone, K. Åström, M. Oskarsson, C. Gustafson, F. Tufvesson
2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), Montreal, QC, 2017, pp. 1-7.

Publications not included in this thesis:

Quality of Academic Writing for Engineering Students at Lund University

I Reinhold, K Batstone, I Gallardo Gonzalez, A Troian, R Yu
The 2nd EuroSoTL conference, June 8-9 2017, Lund, Sweden

Chapter 1

Computer Vision without Vision: Methods and Applications of Radio and Audio Based SLAM

1 Introduction

Navigation has been fundamental to human civilisation and animal-kind alike. By being able to find and return to places rich in resources gives a person or animal a large evolutionary advantage. For humans the principle for navigation has changed very little over time. We have almost always used reference points to navigate relatively to them, ie. Landscape and Stars. With the advancement in knowledge, cartography has played a large part in navigating with more precision, allowing for faster trade routes and military advantages. Now in the modern era, satellites for a Global Positioning System (GPS) are used from mobile devices to vehicles, with a precision of roughly two metres. Although this is a very good precision, issues arise when in urban or precipitous regions due to the reflections and attenuations of the radio waves sent to and from the GPS satellites. This problem then gives rise to the research in this thesis. Due to requirements of modern navigation systems demanding higher precision indoors and in urban areas, robust solutions must be found. Some solutions do exist currently but each have their own drawbacks. One such solution is Bluetooth beacons. They are very cheap and most mobile platforms already have the required existing hardware but they have a short range. For full coverage of an office building, it would require hundreds, if not thousands of beacons to provide a good precision in positioning, with each of the beacons to be calibrated beforehand. For most companies this would be unfeasible. Another such technology could be ultrasound. Most office buildings already have loudspeakers located throughout the building, but young children and animals can hear these frequencies, which for the case of guide dogs can be problematic.

The key challenges in this area, is the ability to calibrate the locations of the broadcasting media, ie. the reference points, to find a suitable media such that the infrastructure already exists and widely available in urban areas and lastly, find a robust solution to calibrate the locations of the broadcasting media so that it can be done though crowdsourced data.

From existing research completed here at the Centre of Mathematical Sciences, LTH, we have solved algebraically, how to find the locations of the broadcasting media and relative positions to them with the smallest amount of measurements required [1, 2]. This is what we call a minimal solver. Further research has also been done on using these minimal solvers to calibrate larger systems using both sound and Wi-Fi signal strength, [3].

In this thesis, the idea of localisation and mapping are explored through the techniques more commonly used in Computer Vision. We have investigated existing and new technologies, along with robust algorithms to simultaneously calibrate the locations of access points and microphone and the relative position of the user. We also look into addressing large volume of crowdsourced data and using this to solve the problem.

2 Time of Arrival

Time of Arrival is a simple method to discover the location of a target. Using three reference points, if you measure the time it takes for a emission, such as a sound event, to go from the target to the reference points, or visa verse, then it is possible to find the targets location. This is computed based on the prior knowledge that, in this case sound, has a constant velocity. Due to the constant velocity, c, time measurements can be converted into distances to each of the three reference points, d_i,

$$d_i = c(t_{i,end} - t_{start}). \tag{1.1}$$

In a 2D space, these distances then constitute as a solution set for each of the reference points, $\mathbf{r}_i$, in the form of circles. In the case of having three or more circle solution sets, then there exist one point that lays on each of these circles. This is the intersection of all circles hence this is the target location, $\mathbf{s}$, (see Figure 1.1). This can be formally expressed as,

$$d_i = \|\mathbf{r}_i - \mathbf{s}\|_2. \tag{1.2}$$

These principles form the basis of the majority of navigation systems, for mobile phone mast triangulation and GPS from satellites with a known position. Here we can see why precision of knowing the reference points and the distance measurements are key to reliable navigation. For satellites, they use atomic clocks to measure the time it takes for a radio signal to reach the user.

What if the reference points are not known? This is the Self Calibration problem. The idea of the self calibration problem is to find the locations of the reference points and the target location at the same time. Based on the existing framework, it can be seen that there will not be one solution. Each reference point will have the freedom to rotate in a circle at distance d_i, around the target location. Further more, the coordinate system will be lost, so a relative coordinate system can only be used (see Figure 1.2).

To constrain this system, more information will be required. One method to do this, is to allow the reference point to communicate with each other and take measurements. This means that the reference points will know the distances between them as well as the distances to the target. This then would form a rigid graph, where all the points can be calculated through Multidimensional Scaling, [4]. This method works very well for ideal measurements, but when this is applied to a realistic environment, it is not always feasible for the reference points to communicate with each other.

Increasing the number of reference points would also not help, once again all the reference points will have the freedom to rotate. One possibility would be to increase the number

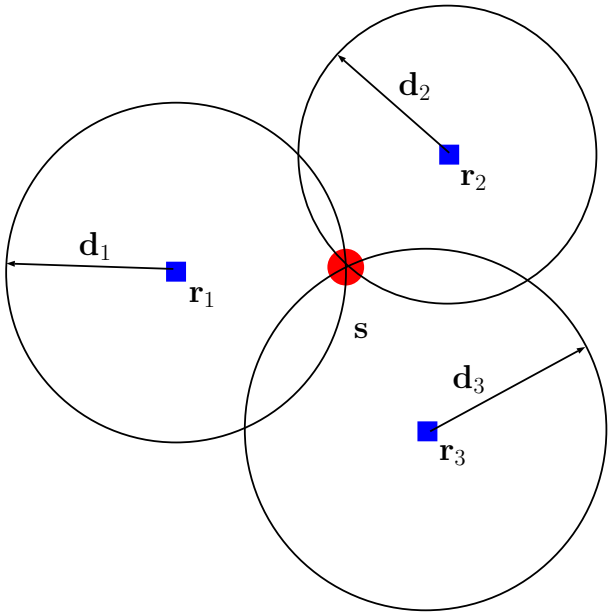


Figure 1.1: Illustration of trilateration. Receivers positions, $\mathbf{r}$, (Blue squares), Sender positions, $\mathbf{s}$, (Red circles) and measurements, $\mathbf{d}$, (black circles and lines) .

of targets. From now on reference points will be called receiver positions and the targets will be called sender positions. If we increase the number of sender positions by two, then there will be enough information, the reason why will be explained in a later section. All senders will then communicate to all receivers for, in this case, a total of nine measurements. This forms a rigid bipartite graph, ie. the receivers have not communicated with any other receiver and the senders have not communicated with any other sender. By having these extra senders, the receivers will no longer be able to rotate around a given sender, as this will break the constraints of the other senders, $\mathbf{s_j}$. This then means that despite not knowing the locations of the receivers, $\mathbf{r_i}$, they must be in a specific configuration in order to satisfy the measurements,

$$d_{ij} = \|\mathbf{r}_i - \mathbf{s}_j\|_2. \tag{1.3}$$

From Algebraic Geometry it is found to have a finite number of solutions/ configurations (see Figure 1.2), [1]. This is discussed further in Section 4.

3 Technologies and Hardware

As discussed before there are currently many systems available for navigation and localisation. This section will go over the different measurement media used in our methods and

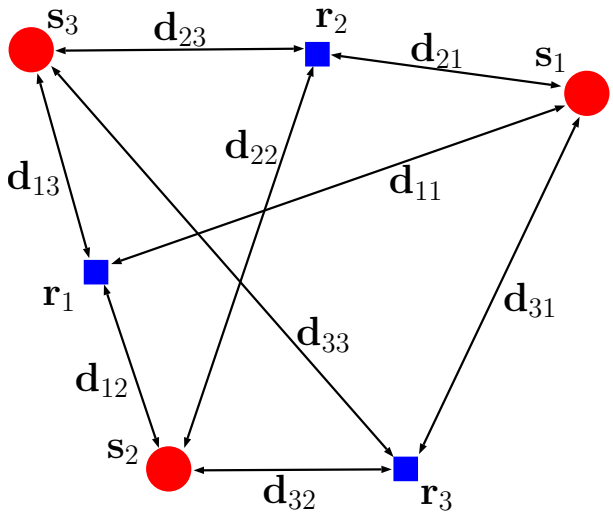


Figure 1.2: Illustration of Self-Calibration of a minimal solution of three receivers and three senders in 2D. Receivers positions, $\mathbf{r}$, (Blue squares), Sender positions, $\mathbf{s}$, (Red circles) and measurements, $\mathbf{d}$, (black lines).

give an overview of the other systems which serve as a comparison.

3.1 Global Positioning System (GPS)

GPS is a Satellite based navigation system. It was first developed by the U.S. Department of Defence in 1978 but became publicly available in 2000. GPS is a radio based system that uses satellites to broadcast their current time and position constantly. In doing so the user on the ground can find the time delay between the time the signal was sent and received. Since the speed of a radio signal is constant, then the distance between the user and the satellite is known. Therefore if the distance to three other satellites are also found, and since the location of the satellites is also known, then it is possible to find your location, this is a trilateration method [5]. Of course the locations calculated using this system are not perfect. The satellites must have atomic clocks in order to have the most stable and accurate time to broadcast as well, but these clocks do drift so there are routines to synchronise these clocks for better performance. Similarly for the GPS device on the Earth. With these methods in place GPS can have a precision of $4.9m$ on average using smartphones, [6]. Further errors arise when the user is in urban areas. The radio signal for the satellites are attenuated and reflected by building and trees. This means that the timed signal is not received or has a multipath component hence the signal is received by the user at a later than normal time, which can lead to large errors. Today GPS is more commonly known as Global Navigation Satellite System (GNSS), which encompasses multiple Satellite systems, GPS, GLONASS, Galileo and BeiDuo. These are used to give a reliable service throughout the globe.

3.2 Received Signal Strength Indicator (RSSI)

RSSI is, more commonly, a Wi-Fi based estimate of the signal strength received at the user from a Wi-Fi router. As a user gets closer to the Wi-Fi router Signal Strength Indicator increases and is measured in terms of *dBm*. The power of the broadcasted radio signal from Wi-Fi router cannot be known to the user whom receives the signal. In order to estimate the distance between the user and the Wi-Fi router, a path loss model is commonly used [7]. An example of such a model is

$$P = C - 10\gamma log_{10}(d) + X, \tag{1.4}$$

where P is the RSSI value in *dBm*, C is the measured power at $1m$, γ is the path loss exponent, d is the estimated distance and X is a normally distributed noise. More formally, it can be seen that to calculate the distance, C and γ must be estimated to give an accurate distance. These types of measurements can be used in Time-Difference-of-Arrival (TDOA) systems, since the unknown terms create offsets between the user and Wi-Fi router so the difference in relative RSSI are used to estimate position.

RSSI measurements unfortunately suffer from many multipath components, which can make it difficult to have reliable localisation systems. Most Wi-Fi routers are located indoors where environments are complex and have many walls, which can attenuate the radio signals.

3.3 Ultra-Wideband (UWB)

Ultra-Wideband devices are commercially available, radio based low powered devices that broadcast on a large bandwidth. More interestingly to us it also comes with a protocol, which allows UWB devices to range between devices using two-way timing, [8]. Due to its low energy consumption, these devices are ideal for robots and Internet of Things. In our work we have used the Decawave DWM1000 chip on small quadcopter drones, as seen in Figure 1.3.

These UWB device are becoming more common place, since September 2019 Iphones have had UWB capabilities and other major companies are looking to follow. Unlike RSSI, the two-way timing protocol allows for more reliable measurements. Two-way timing works by the sender device sending packets back and forth with the receiver UWB device. The sender first sends a Range request at time t_{s_1} to the receiver at time t_{r_1} with its ID. The receiver then responds after a delay at time t_{r_2} which then arrives at the sender again at time t_{s_2}. After another short delay at time t_{s_3}, the sender then communicates the ID, t_{s_1}, t_{s_2} and t_{s_3} back to the receiver, see Figure 1.4.

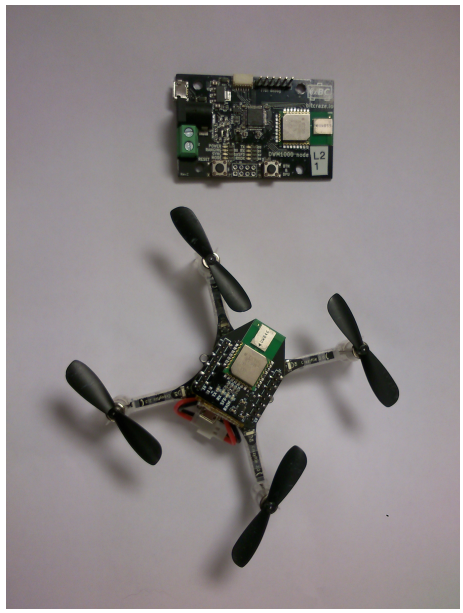


Figure 1.3: The Bitcraze quadcopter with Decawave DWM1000 UWB chips.

The receiver can then calculate the range, d, using the data it has received by the using,

$$d = \frac{c((t_{s_2} - t_{s_1}) - (t_{r_2} - t_{r_1}) + (t_{r_3} - t_{r_2}) - (t_{s_3} - t_{s_2}))}{4}, \tag{1.5}$$

for a constant c, the speed of light. At this point the receiver can then report back to the sender the measured distance, but this is optional.

For this thesis effort was made to characterise these measurements, to try find a suitable model that would describe them. During the writing of this thesis, there were publications on characterising these measurements, such as [9], but many used custom antennas, which you would not find on an commercial device, as it may skew the reported precision of such devices. Therefore an experiment was conducted to find these characteristics using the Bitcraze quadcopter and the Decawave DWM1000 UWB devices.

As it can bee seen in Figure 1.5, there is a distribution with an overall mean of $4.6073m$ and a standard deviation of $0.1214m$. This distribution appears to be a combination of two Gaussian distributions. The main distribution occurs around $4.5m$ and another appears to be at around $4.8m$. The second distribution is believed to be a reflection off the table that the receiver with a stand was on. The stand that held receiver was $12cm$ tall, which would correspond to the second distribution. Looking again at the first distribution, the measurements do correspond to a Gaussian distribution with an error of $\pm 0.2m$.

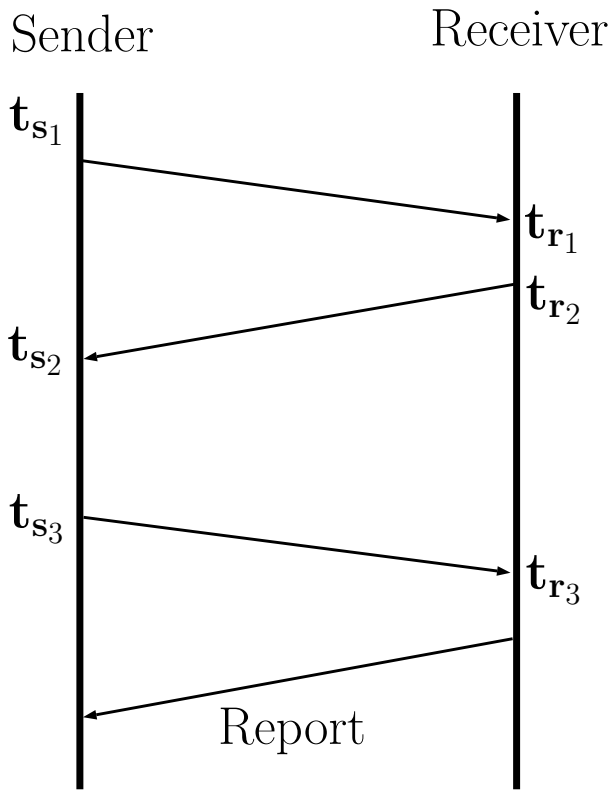


Figure 1.4: Illustration of Two-way timing protocol.

3.4 Wi-Fi 802.11mc Round-Trip Time (RTT)

The IEEE Standard 802.11mc, [10], is a Wi-Fi protocol for performing Round-Trip Time measurements. It is also known as Fine Time Measurements (FTM). This protocol is relatively new, although the hardware for this protocol has been around for years, it has only become widely commercially available for the past year. With the release of Android 10, this protocol as seen more widespread use, with a handful of routers and mobile phones that currently support it. Similarly to UWB this form of RTT measurements achieves higher precision than RSSI due to less interference from multipath components. RTT works by the sender (mobile phone) device sending packets back and forth with the receiver (router). The sender first sends a Range request. The receiver then responds after a delay at time t_{r_1} which then arrives at the sender at time t_{s_1}. After a short delay at time t_{s_2} the sender then communicates back to the receiver arriving at t_{r_2}. The times t_{r_1} and t_{r_2} are then sent back to the sender for the distance to be calculated, see Figure 1.6.

Similarly to before the distance, d, is calculated using,

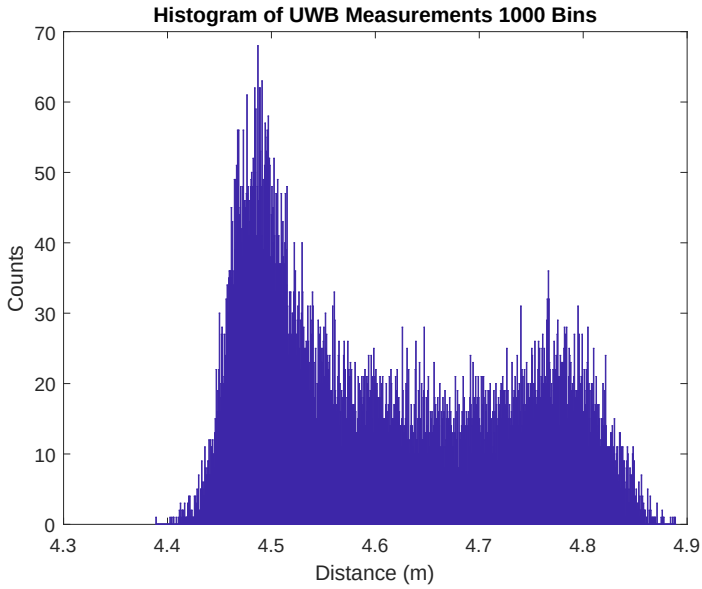


Figure 1.5: Histogram of 17647 UWB Two-way time measurements at a fixed distance of $4.55m$.

$$d = \frac{c((t_{r_2} - t_{r_1}) - (t_{s_2} - t_{s_1}))}{2}, \tag{1.6}$$

for a constant c, the speed of light.

Unlike UWB two-way timing, when this protocol is used in Android devices, this RTT measurement is repeated in bursts to give a more accurate measurement, but also a standard deviation is returned to the sender giving additional measurements to the user. Once again, due to this protocol being relatively new, experiments were performed to find the characteristics of the measurements. A Google Pixel 4 XL mobile phone was used with Android 10. The Google Mesh routers were used as the receivers.

As it can be seen in Figure 1.7, once again we have a Gaussian distribution with a mean of $4.75m$ and a standard deviation of $0.33m$. It can be noted that there appears to be an offset from the groundtruth distance $4.19m$. This offset has only been noticed when using the Android implementation and not previous implementations used for publications. This offset is mentioned in the notes for the RTT API in Android but regardless of this, the precision of RTT is usually $\pm 1.2m$.

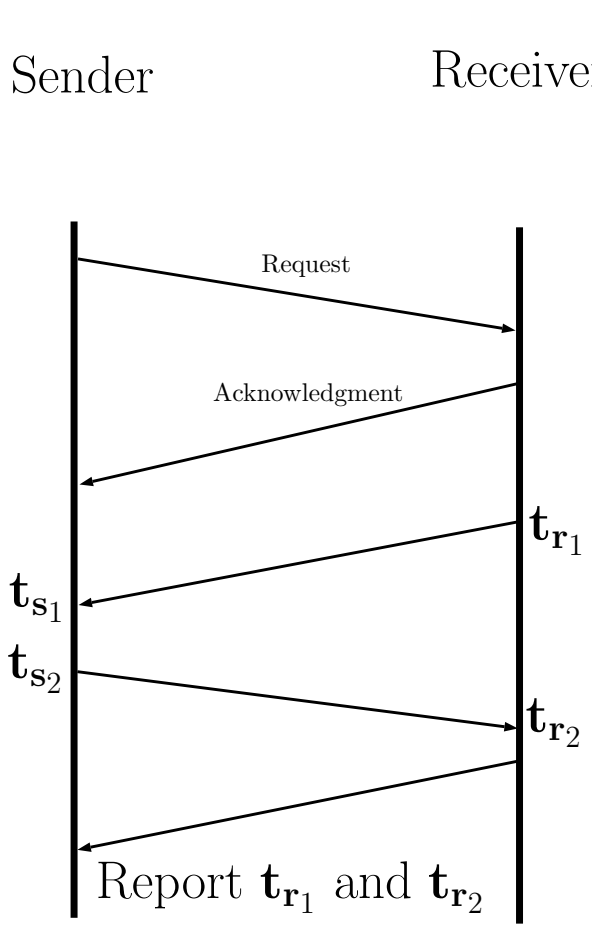


Figure 1.6: Illustration of RTT protocol.

3.5 Massive Multiple Input and Multiple Output (Massive MIMO)

Massive MIMO is a relatively new technology, the main principle of which is to have base stations with a large number of antennas. In doing so these types of base stations have the ability to provide a good service to many terminals at the same time. This technology is currently being used to provide 5G services to many users at high bandwidths. So what use is this to positioning? By having many antennas in an array, it is possible to obtain the Angle of arrival (AoA) and Angle of Departure (AoD) of a communicating signal in addition to RSSI. Further to this it is possible to find out if the signal was Line of Sight or not, giving further information on the geometry of the environment, [11].

With regards to the accuracy of the measurements, this is still relatively unknown. In paper VII, we estimate a accuracy of 13 cm. There has been many different novel techniques, but no standardisation. The aim from this technology is to improve the accuracy from LTE

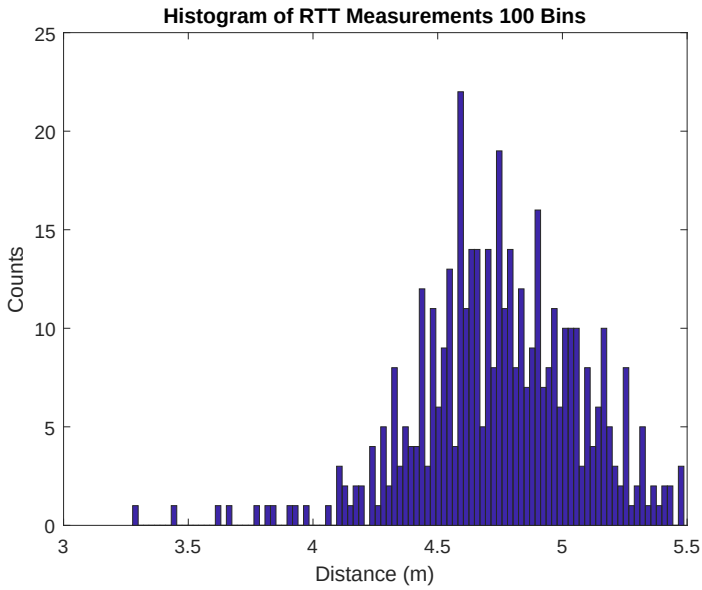


Figure 1.7: Histogram of 2987 RTT measurements at a fixed distance of $4.19m$.

which has an accuracy of $10m$, [12].

4 Time-of-Arrival Minimal Solvers

As discussed in Section 2, Time-of-Arrival (TOA) can have a fixed number of solutions for the self-calibration problem. In order to find the receiver positions $\mathbf{r}_i$ and sender positions $\mathbf{s}_j$, we can look at the square form of Equation (1.3),

$$d_{ij}^2 = \|\mathbf{r}_i - \mathbf{s}_j\|_2^2, \tag{1.7}$$

$$\Rightarrow d_{ij}^2 = (\mathbf{r}_i - \mathbf{s}_j)^T(\mathbf{r}_i - \mathbf{s}_j), \tag{1.8}$$

$$\Rightarrow d_{ij}^2 = \mathbf{r}_i^T\mathbf{r}_i + \mathbf{s}_j^T\mathbf{s}_j - 2\mathbf{r}_i^T\mathbf{s}_j. \tag{1.9}$$

This can be thought of as equivalent to the original TOA problem equation (1.3) since the distance measurements are real and positive. By performing a series of invertible linear combinations of d_{ij}^2 we can form

$$\mathbf{B} = \begin{pmatrix} d_{11}^2 & d_{12}^2 - d_{11}^2 & \dots & d_{1n}^2 - d_{11}^2 \\ d_{21}^2 - d_{11}^2 & & & \\ \dots & & \hat{\mathbf{B}} & \\ d_{m2}^2 - d_{11}^2 & & & \end{pmatrix}, \tag{1.10}$$

where the compaction matrix $\hat{\mathbf{B}} \in \mathbb{R}^{(m-1)\times(n-1)}$ can be defined as having entries $\hat{B}_{ij} = \frac{d_{i,j}^2 - d_{i1}^2 - d_{1j}^2 + d_{11}^2}{-2}$, with $i = 2, \ldots m$ and $j = 2, \ldots, n$. The first row and column of the matrix $\hat{\mathbf{B}}$ are used as constraints for the solution.

Here we can form a factorisation problem for $\hat{\mathbf{B}}$, where $\hat{\mathbf{B}} = \mathbf{R}^T\mathbf{S}$. The matrix $\mathbf{R}_i$ can be thought of as the vector from the first receiver $\mathbf{r}_1$ to the receivers $\mathbf{r}_i$, *i.e.* $\mathbf{R}_i = \big[(\mathbf{r}_i - \mathbf{r}_1)\big]$, similarly for the matrix $\mathbf{S}_j = \big[(\mathbf{s}_j - \mathbf{s}_1)\big]$. This formulation of $\mathbf{R}$ and $\mathbf{S}$ then implies that they must be in a 3D affine space, which in turn implies that the matrix $\hat{\mathbf{B}}$ has at most rank 3. Due to this fact, we can use this information to make a low rank approximation of the matrix, this is explained further in section 6.1. This rank constraint also implies that we require $m \geq 4$ receivers and $n \geq 4$ senders.

By fixing $\mathbf{r}_1 = 0$ at the origin and $\mathbf{s}_1 = \mathbf{L}\mathbf{b}$ as a vector from the origin for an invertible transformation matrix $\mathbf{L}$ and vector $\mathbf{b}$. Hence, the problem is reformulated as follows,

$$\begin{aligned} \mathbf{r}_i &= \mathbf{L}^{-T}\tilde{\mathbf{R}}_i,\ i = 2 \ldots m, \\ \mathbf{s}_j &= \mathbf{L}(\tilde{\mathbf{S}}_j + \mathbf{b}),\ j = 2 \ldots n, \end{aligned} \tag{1.11}$$

where $\tilde{\mathbf{R}} = \mathbf{L}^T\mathbf{R}$, $\tilde{\mathbf{S}} = \mathbf{L}^{-1}\mathbf{S}$, and hence $\hat{\mathbf{B}} = \tilde{\mathbf{R}}^T\mathbf{L}^{-1}\mathbf{L}\tilde{\mathbf{S}} = \mathbf{R}^T\mathbf{S}$.

Using this parametrisation, the constraints from the first row and columns of matrix $\mathbf{B}$, become

$$\begin{aligned} d_{11}^2 &= (\mathbf{r}_1 - \mathbf{s}_1)^T(\mathbf{r}_1 - \mathbf{s}_1) = \mathbf{s}_1^T\mathbf{s}_1 = \mathbf{b}^T\mathbf{L}^T\mathbf{L}\mathbf{b} \\ &= \mathbf{b}^T\mathbf{H}^{-1}\mathbf{b}, && (1.12) \\ d_{1j}^2 - d_{11}^2 &= \mathbf{s}_j^T\mathbf{s}_j - \mathbf{s}_1^T\mathbf{s}_1 = \tilde{\mathbf{S}}_j^T\mathbf{L}^T\mathbf{L}\tilde{\mathbf{S}}_j + 2\mathbf{b}^T\mathbf{L}^T\mathbf{L}\tilde{\mathbf{S}}_j \\ &= \tilde{\mathbf{S}}_j^T\mathbf{H}^{-1}\tilde{\mathbf{S}}_j + 2\mathbf{b}^T\mathbf{H}^{-1}\tilde{\mathbf{S}}_j, && (1.13) \\ d_{i1}^2 - d_{11}^2 &= \mathbf{r}_i^T\mathbf{r}_i - 2\mathbf{r}_i^T\mathbf{s}_1 = \tilde{\mathbf{R}}_i^T(\mathbf{L}^T\mathbf{L})^{-1}\tilde{\mathbf{R}}_i - 2\mathbf{b}^T\tilde{\mathbf{R}}_i \\ &= \tilde{\mathbf{R}}_i^T\mathbf{H}\tilde{\mathbf{R}}_i - 2\mathbf{b}^T\tilde{\mathbf{R}}_i, && (1.14) \end{aligned}$$

where the symmetric matrix $\mathbf{H} = (\mathbf{L}^T\mathbf{L})^{-1}$. In order to solve this system of equations, there are in total nine unknowns, six unknowns for $\mathbf{L}$ and three unknowns for $\mathbf{b}$.

The symmetric matrix $\mathbf{H} = (\mathbf{L}^T\mathbf{L})^{-1}$ can also be written as $\mathbf{H} = \frac{adj\mathbf{H}}{det\mathbf{H}}$, hence the system of equations become,

$$\begin{aligned} d_{11}^2 det\mathbf{H} - \mathbf{b}^T adj\mathbf{H}\mathbf{b} &= 0, && (1.15) \\ (d_{1j}^2 - d_{11}^2)det\mathbf{H} - \tilde{\mathbf{S}}_j^T adj\mathbf{H}\tilde{\mathbf{S}}_j - 2\mathbf{b}^T adj\mathbf{H}\tilde{\mathbf{S}}_j &= 0, && (1.16) \\ d_{i1}^2 - d_{11}^2 - \tilde{\mathbf{R}}_i^T\mathbf{H}\tilde{\mathbf{R}}_i + 2\mathbf{b}^T\tilde{\mathbf{R}}_i &= 0. && (1.17) \end{aligned}$$

This system of equations can be solved as a system of polynomials. As we can see from equation (1.17), this equation is linear in its constraints, whereas equations (1.15) and (1.16) are not. We can therefore turn to a branch of mathematics called Algebraic Geometry. Algebraic Geometry classically is the study of solving multivariate polynomials, but more recently has also introduced computational methods and software packages, such as Macaulay2 [13], for solving polynomials algebraically.

For ease of understanding, I will introduce very basic concepts of Algebraic Geometry. Further details of which can be found at [14]. We can define a polynomial f as a finite linear combination of monomials $X = \{x_1, x_2, \ldots, x_n\}$ which is a finite product of variables. With a set of polynomials, we can form what is known as an Ideal I. An Ideal can be formulation as follows,

$$I = \langle f_1, f_2, \ldots, f_m \rangle = \Big\{ \sum_{l=1,m} h_l f_l \mid h_l \in \mathbb{C}[X] \Big\}. \tag{1.18}$$

Since all the polynomials $f = 0$ for a some selection X then the polynomials can also be expressed as a linear combination. The solutions to this Ideal I, is known as a variety $V(I)$, where

$$V(I) = \{x \in \mathbb{C}^n \mid f(x) = 0, \forall f \in I\}. \tag{1.19}$$

This is essentially the goal for polynomial solving. We wish to find the values for the monomials such that the polynomials are solved. Properties that are found to be advantageous, is that a Gröbner Basis G can be formed. A Gröbner basis is a type of generating set of polynomials for an ideal I. In simpler terms this is a set basis polynomials that can be used to form the polynomials in an Ideal. The advantage of the Gröbner Basis is that the number of possible monomials less than highest order terms in the Ideal is the same as the number of solutions in the variety $V(I)$. There are various methods to finding such Gröbner Bases and Varieties, details of which can be found [14, 15].

Going back to the problem at hand, equations (1.15),(1.16) and (1.17) can be used to form an ideal. For the $m = n = 5$ minimal solver, the linear constraints in (1.17) reduce the system to five equations in (1.15) and (1.16) with five unknowns $X = \{x_1, x_2, x_3, x_4, x_5\}$ and 42 solutions for $det(\mathbf{H}) \neq 0$, [2].

5 Random Sample Consensus

Random sample consensus (RANSAC) is a key tool for robust model estimation. The algorithm was first published by Fischler and Bolles as a method for the Local Determination Problem, [16]. When using data that has large amounts of errors, this can cause havoc for optimization solvers leading to incorrect local minima solutions or divergence. One way

to work around this, is to try and determine the data that has the errors, in other words outlier detection. In doing so, you can find a subset of the data that does correspond to a given model. The RANSAC method takes advantage of minimal solvers. A minimal solver uses the smallest amount of data in order to obtain a solution. These can be computed quickly and give a solution which can then be analysed. RANSAC can be broken down into 5 main steps.

1. Randomly select a minimal subset of the data required for a model
2. With this subset compute the model parameters
3. Using a predefined threshold, find how many other points fit this model
4. Repeat steps 1-3, N times
5. Choose the model with the most number of points and best fit. This will be the inlier set.

The simplest example of the RANSAC method is 2D line fitting. To fit a line, only two data points are required. As seen in Figure 1.8, there are two possible line fittings have been drawn l_1 and l_2. One of which is clearly the better solution. By using a threshold, ϵ, then we can quantify which of the two lines is better, by counting the number of points that are in the region of the threshold.

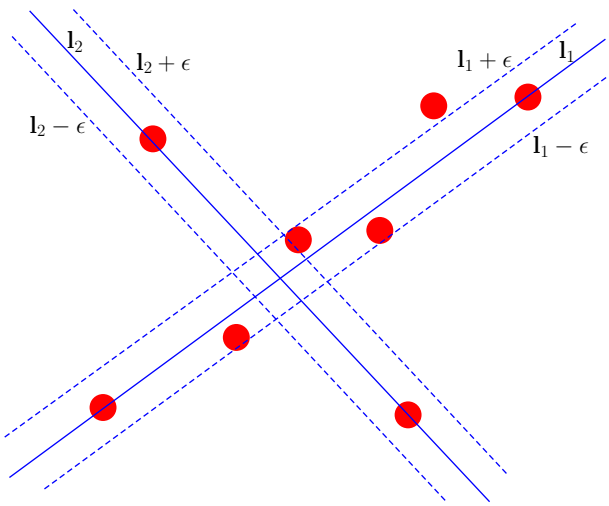


Figure 1.8: Illustration of RANSAC Line fitting. Two possible lines (blue) are drawn with their thresholds (dashed blue) and data points (red).

By iterating enough times, it can bee seen that a solution can be found quickly and the estimate of the inliers can be found. Now when calculating the final model parameters, we

can use the inlier dataset and ignore the gross errors to hopefully give the global optimum. This algorithm has two main issues. One issue is trying to determine the threshold used in step 3. Too large and there is a possibility of introducing too many points with gross error, too little then there is a possibility of not finding a good estimate of the inlier set. The other issue, is to determine the number of iterations, too few and the correct inliers set can be missed, too many is a waste of computational time, for very little gain. Regardless of these issues, RANSAC itself does not guarantee optimality, [17], but there are variants of the RANSAC algorithm that does, [18].

6 Matrix Factorisation

For many computer vision problems, such as structure-from-motion [19], matrix factorisation is a key part in the calculation. Given a matrix $\mathbf{H} \in \mathbb{R}^{m \times n}$ of rank r, this can be factorised as

$$\mathbf{H} = \mathbf{AB}, \tag{1.20}$$

where $\mathbf{A} \in \mathbb{R}^{m \times r}$ and $\mathbf{B} \in \mathbb{R}^{r \times n}$. This factorisation can be performed in many different ways, using different algorithms which can give different results. The different results are fine, as there is no unique solution. For every solution there exists an invertible matrix $\mathbf{L} \in \mathbb{R}^{r \times r}$ such that

$$\mathbf{H} = \mathbf{AB} = (\mathbf{AL})(\mathbf{L}^{-1}\mathbf{B}). \tag{1.21}$$

The rank of a matrix $\mathbf{H}$ is the maximum number of linearly independent columns. For matrix $\mathbf{H}$, this means

$$\text{rank}(\mathbf{H}) \leq \min(m, n). \tag{1.22}$$

With rank in mind a common factorisation method is Singular Value Decomposition (SVD), which is denoted as follows,

$$\mathbf{H} = \mathbf{USV}^T \tag{1.23}$$

where $\mathbf{U} \in \mathbb{R}^{m \times m}$ and is orthogonal, $\mathbf{U} \in \mathbb{R}^{n \times n}$ and is orthogonal and $\mathbf{S} \in \mathbb{R}^{m \times n}$ and is a diagonal matrix with non-negative elements. The elements of $\mathbf{S}$ are the singular values $\sigma_i(\mathbf{H})$ and are commonly ordered in descending order, ie.

$$\sigma_1(\mathbf{H}) \geq \sigma_2(\mathbf{H}) \geq ... \geq \sigma_r(\mathbf{H}) \geq 0. \tag{1.24}$$

Here the rank is equal to the number of non-zero singular values.

6.1 Low Rank Matrix Approximation

In this thesis many of the problems have full rank, due to using measurements which have errors. From studying the problem we know that these problems should have a fixed rank.

Therefore it is possible to approximate the measurements with a fixed rank. The rank approximation can be formulated as follows, [20],

$$\min_{\mathbf{K}} \|\mathbf{H} - \mathbf{K}\|_F^2 \text{ subject to } \text{rank}(\mathbf{K}) = t \leq r, \tag{1.25}$$

where $\mathbf{K}$ is a low rank approximation of $\mathbf{H}$. This can be found by truncating the singular matrix $\mathbf{S}$ in $\mathbf{H} = \mathbf{U}\mathbf{S}\mathbf{V}^T$, ie.

$$\mathbf{H} = \mathbf{U}\text{diag}(\sigma_1, \sigma_2, ..., \sigma_r)\mathbf{V}^T \approxeq \mathbf{U}\text{diag}(\sigma_1, \sigma_2, ..., \sigma_t, 0, ..., 0)\mathbf{V}^T = \mathbf{K}. \tag{1.26}$$

Now for a problem where the rank is known, for example 2, then

$$\hat{\mathbf{H}} = \mathbf{U}\text{diag}(\sigma_1, \sigma_2, 0, ..., 0)\mathbf{V}^T. \tag{1.27}$$

6.2 Matrix Factorisation with Missing Information

Again when working with real measurements it is never guaranteed the that you get measurements all the time. When working with Wi-Fi for instance, walls can attenuate the radio signal or in computer vision image points can be obscured by other object. Due to these reasons, the matrix you wish to factorise has missing data. Therefore strategies are created to work around these problem areas. One such strategy is to reorder the matrix in such a way, so that all the missing data is in one smaller block of the matrix and a full block is created which can then be factorised.

For large amounts of missing data, such as structure for motion problems, reorganising the matrix is not possible. These types of problems often form a banded matrix structure. This is due to the points of an object being obscured by the object itself as the camera rotates around the object. For these types of problems most of the matrix has missing data. The strategy for this problem is to solve sub-blocks of the matrix. We can define a matrix $\mathbf{H}$ with rank r, which consists of two sub-matrices $\mathbf{H}_A$ and $\mathbf{H}_B$ where $\text{rank}(\mathbf{H}) = \text{rank}(\mathbf{H}_A) = \text{rank}(\mathbf{H}_B) = r$, as shown in Figure 1.9. Both sub-matrices contain no missing entries. This in turn allows for each sub-matrix to be factorised,

$$\mathbf{H}_A = \mathbf{U}_A\mathbf{V}_A^T, \tag{1.28}$$
$$\mathbf{H}_B = \mathbf{U}_B\mathbf{V}_B^T. \tag{1.29}$$

Since the factorisations are not unique then both $\mathbf{H}_A$ and $\mathbf{H}_B$ have ambiguities, that is there exists an invertible matrix $\mathbf{L}$ such that,

$$\mathbf{H}_B = (\mathbf{U}_B\mathbf{L})(\mathbf{L}^{-1}\mathbf{V}_B^T) \tag{1.30}$$

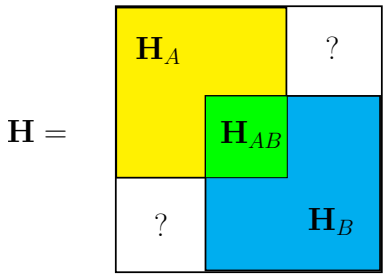


Figure 1.9: Illustration of the matrix $\mathbf{H}$ with it two sub-blocks $\mathbf{H}_A$ (yellow) and $\mathbf{H}_B$ (blue). The overlapping region, $\mathbf{H}_{AB}$, is shown in green and the missing data, **?**, is in white.

In this current form, we have two separate factorisations, if these two sub-matrices $\mathbf{H}_A$ and $\mathbf{H}_B$ have a sufficiently large overlapping region, which we can define as $\mathbf{H}_{AB} \in \mathbb{R}^{m \times n}$. Here m and n must be $\geq r$ in order to maintain the rank of $\mathbf{H}$, hence a sufficiently large overlapping region. Since the overlapping region exists in both factorisations, then

$$\mathbf{H}_{AB} = \hat{\mathbf{U}}_A \hat{\mathbf{V}}_A^T = \hat{\mathbf{U}}_B \hat{\mathbf{V}}_B^T, \tag{1.31}$$

$$\Rightarrow \mathbf{H}_{AB} = \hat{\mathbf{U}}_A \hat{\mathbf{V}}_A^T = (\hat{\mathbf{U}}_B \mathbf{L})(\mathbf{L}^{-1} \hat{\mathbf{V}}_B^T), \tag{1.32}$$

where $\hat{\mathbf{V}}$ and $\hat{\mathbf{U}}$ are the corresponding sub regions in the factorisation of $\mathbf{H}_A$ and $\mathbf{H}_B$.

Now we can find $\mathbf{L}$ such that

$$\hat{\mathbf{U}}_A = \hat{\mathbf{U}}_B \mathbf{L}, \tag{1.33}$$

$$\hat{\mathbf{V}}_A^T = \mathbf{L}^{-1} \hat{\mathbf{V}}_B^T, \tag{1.34}$$

this can be solve by least squares and forms the optimisation problem,

$$\min_{\mathbf{L}} \|\hat{\mathbf{U}}_A - \hat{\mathbf{U}}_B \mathbf{L}\|_F^2 + \|\mathbf{L} \hat{\mathbf{V}}_A^T - \hat{\mathbf{V}}_B^T\|_F^2. \tag{1.35}$$

Using this solution, one can form the whole factorisation by concatenating $\mathbf{U}_A$ with $\mathbf{U}_B \mathbf{L}$ excluding the overlapping region and similarly with $\mathbf{U}\mathbf{V}_A$ with $\mathbf{L}^{-1}\mathbf{V}_B$ again excluding the overlapping region. Hence

$$\mathbf{H} = \mathbf{U}\mathbf{V}^T = \begin{bmatrix} \mathbf{U}_A \\ \tilde{\mathbf{U}}_B \mathbf{L} \end{bmatrix} \begin{bmatrix} \mathbf{V}_A^T & \mathbf{L}^{-1} \tilde{\mathbf{V}}_B^T \end{bmatrix}, \tag{1.36}$$

where $\tilde{\mathbf{U}}$ and $\tilde{\mathbf{V}}$ represent the non overlapping region.

7 Summary of Estimation Problems

In this section, I will give a summary of the estimation problems studied in this thesis. As discuss in previous sections of this thesis, when the measurements are perfect, the theory already exists to solve problems, like Structure-from-Motion and TOA self calibration. Issues arise when working with real measurements, where not all measurements are accurate and in fact sometimes missing. For these problems, robust algorithms must be made to make this theory useable and reliable. Throughout the research conducted for this thesis, each of the problems must undergo a non-linear optimisation stage to find the local minimum. This form of optimisation is very sensitive to outliers (bad data), and if the outliers are included, we will arrive at a unusable local minimum. In the cases studied as part of this thesis, we test and formulate methods to classify inliers (good data) from the outliers. By optimising over the inlier set, we can form usable solutions.

7.1 Low Rank Matrix Factorisation with Missing Data and Outliers

When it comes to low rank matrix factorisation with missing data and outliers, the previous section has shown how to handle missing data and low rank approximations. When working with problems in computer vision geometry, such as Affine Structure-from-Motion problems, where the observation matrix is factorised to obtain the camera motion and the 3D structure, these types of problems have structured data but do contain missing data and outliers. To overcome the outliers, an alternative formulation of matrix factorisation must be made. This section forms the basis for Paper IV and has influence on Paper V. Given the matrix $\mathbf{H} \in \mathbb{R}^{m \times n}$ and we wish to find a low rank factorisation given by the SVD, $\mathbf{K} = \hat{\mathbf{U}}\mathbf{S}\hat{\mathbf{V}}^T$ where $\text{rank}(\mathbf{K}) = t$. Here we can shorten $\mathbf{K} \in \mathbb{R}^{m \times n}$ by combining the singular matrix $\mathbf{S}$ with one of the unitary matrices $\hat{\mathbf{U}}$ or $\hat{\mathbf{V}}^T$, such that $\mathbf{K} = \mathbf{U}\mathbf{V}^T$, where $\mathbf{U} \in \mathbb{R}^{m \times t}$ and $\mathbf{V}^T \in \mathbb{R}^{t \times n}$. Then is it possible to formulate the problem as,

$$e = \|\mathbf{H} - \mathbf{U}\mathbf{V}^T\|_F. \tag{1.37}$$

Since the matrix $\mathbf{H}$ contains missing elements, we can introduce an indexing matrix W with dimensions $m \times n$ whose elements are

$$W_{i,j} = \begin{cases} 1 & \text{if } \mathbf{H}_{i,j} \in \mathbb{R}, \\ 0 & \text{otherwise.} \end{cases} \tag{1.38}$$

. Therefore the formulated problem can be amended as follows,

$$e = \|(\mathbf{H} - \mathbf{U}\mathbf{V}^T) \odot W\|_F, \tag{1.39}$$

where $\odot$ represents element-wise multiplication and a residual can be formed ar $r_{i,j} = (\mathbf{H}_{i,j} - \mathbf{U}_{i,\cdot}\mathbf{V}^T_{\cdot,j}) \odot W_{i,j}$. From understanding the type of data we are working with, it

can be assumed that inlier residuals approximately follow a Gaussian distribution, whereas outlier residuals have approximately uniformly distributed errors. This assumption in turn makes it better to form a truncated squared error.

$$l(r_{i,j}) = \begin{cases} r_{i,j}^2 & \text{if } |r_{i,j}| \leq \epsilon, \\ \epsilon^2 & \text{otherwise.} \end{cases} \tag{1.40}$$

For some reasonable error limit ϵ. With the assumption of good measurements, the problem formulation for low rank matrix factorisation with missing data and outliers can then be made as follows,

$$\min_{\mathbf{U},\mathbf{V}} \quad \sum_{i,j} l(r_{i,j}). \tag{1.41}$$

It is further possible to update the indexing matrix W, to only include the inlier set, hence outliers and missing data can be set to zero.

7.2 Time-of-Arrival Estimation with Missing Data and Outliers

This section forms the basis for Papers I, II, V and VI. In Section 4, it has been discussed how to solve TOA problems. Once again, these solvers rely on accurate measurements, which is not possible when using a media, such as WiFi IEEE 802.11mc, that have a variety of erroneous measurement due to complex indoor environments. Much of this thesis focuses on creating robust methods for solving such issues. By studying the characteristics of the TOA measurements for a particular media, we can estimate the type of noise these measurements have. For the most part of thesis, inlier measurement typically have a Gaussian distribution, therefore the problem can be formulated as follows,

$$d_{ij} = \|\mathbf{r}_i - \mathbf{s}_j\|_2 + \epsilon_{ij}. \tag{1.42}$$

Here the errors $\epsilon_{ij} \in N(0, \sigma)$ are assumed to have a zero mean Gaussian distribution. In reality only inlier measurements will be a zero mean Gaussian distribution, for a complex environment, like an office, radio signals reflect off surfaces causing multipath components, and some measurements are missed completely. With this in mind, we can then reformulate the problem to function for only inlier measurements, giving rise to the optimisation problem,

$$\min_{\mathbf{r},\mathbf{s}} \sum_{(i,j)\in\tilde{W}} (d_{i,j} - ||\mathbf{r}_i - \mathbf{s}_j||_2)^2. \tag{1.43}$$

Here, once again we introduce an indexing matrix W with dimensions $m \times n$ whose elements are

$$W_{i,j} = \begin{cases} 1 & \text{if } d_{i,j} \in \mathbb{R}, \\ 0 & \text{otherwise.} \end{cases} \tag{1.44}$$

At this point the problem formulation can ignore the missing information but all the measurement are still assumes to be inliers. If the subset $\tilde{W}$ contains no outliers and if the initialisation for the optimisation problem is good, then the l^2-norm can give good estimates. Otherwise the optimisation will find a poor estimates. It is imperative that we try to exclude the outlier measurements so we can find good estimates of receiver positions $\mathbf{r}$ and sender positions $\mathbf{s}$. Finding good estimates of the receiver positions $\mathbf{r}$, has more importance for real world calibration, as other less computationally expensive methods can be used to track senders but they require good estimates of the receiver positions $\mathbf{r}$. Hence a strategy can be made, such that the focus is finding good measurements.

In our papers we present a RANSAC scheme to identify these inliers. By understanding the geometry of the model used, it can be found that the compaction matrix, $\hat{\mathbf{B}}$, must be of fixed rank with relation to the dimensionality of model, *i.e.* $\text{rank}(\hat{\mathbf{B}}) = 3$ for a 3D setup. Using this fact, a minimal number of receivers are selected along with a minimal number of senders to form a minimal $\hat{\mathbf{B}}$ with a fixed rank. This is then factorised and a unit vector $\mathbf{v}$ is selected from the left null space. Using the same receiver selection, a compaction matrix $\hat{\mathbf{C}}_{test}$ is formed using the rest of the senders, that have not missing data. In doing so we can now test to see if the rest of the data is orthogonal to the selected unit vector $\mathbf{v}$ *i.e.* if a vector from $\hat{\mathbf{C}}_{test}$, which we can define as $\mathbf{c}_{test}$, is orthogonal to $\mathbf{v}$, that can be expressed as $|\mathbf{v}.\mathbf{c}_{test}| \leq T$ where T is a reasonable tolerance. For the measurement that satisfy $|\mathbf{v}.\mathbf{c}_{test}| \leq T$, these measurements can be defined as an inlier. This is then repeated many times to find the most amount of inliers.

When a inlier set has been identified, this can then be used to solve the TOA self calibration problem. The inliers are used to form a complete and low noise compaction matrix $\hat{\mathbf{B}}_{opt}$, and minimal solvers are used to find the parameters of $\mathbf{L}$ and $\mathbf{b}$. This then hopefully creates a good initial estimate of the receiver positions $\mathbf{r}$ and sender positions $\mathbf{s}$. At this stage we can perform Levenberg-Marquardt optimisation, [21, 22], to refine the receiver positions $\mathbf{r}$ and sender positions $\mathbf{s}$ and solving the optimisation problem equation (1.43). From this point, residuals that don't follow the error term ϵ_{ij} and are non Gaussian, can also be classified as an outlier. The process of finding inliers and classifying outliers can be repeated multiple times, to find the most amount of inliers while maintaining good estimates.

7.3 Constant Offset Time-Difference-of-Arrival Estimation with Missing Data and Outliers

This section deals only with Paper III, but the general concepts are seen in all papers. Constant Offset Time-Difference-of-Arrival (COTDOA) self-calibration is the problem of estimating receiver positions $\mathbf{r}$ and sender positions $\mathbf{s}$ in the presence of a unknown constant offset. This problem lays between TDOA and TOA. Time-difference-of-arrival problem, [23], is the problem of estimating receiver positions $\mathbf{r}$ and sender positions $\mathbf{s}$ which can

allow for a different offset o for every j, *i.e.*

$$z_{ij} = \|\mathbf{r}_i - \mathbf{s}_j\|_2 + o_j. \tag{1.45}$$

Time-of-Arrival however has no offset o_j.

The type of media that is suitable for COTDOA would be a repeating signal, such as a repeating sound chirp. The waveform of such a chirp is irrelevant for this scenario, just that it is repeated on a regular period. These chirp emissions occur at time T_j and are unknown, but can be written as follows,

$$T_j = k_1 j + k_0, \tag{1.46}$$

where k_1 is the known interval.

Similar to the other problems like TOA, the problem we are considering involves m receiver positions $\mathbf{r}_i$ and n sender positions $\mathbf{s}_j$. If the chirp event j is then detected at receiver positions $\mathbf{r}_i$ at time t_{ij}, then a distance can then be calculated,

$$c(t_{ij} - T_j) = \|\mathbf{r}_i - \mathbf{s}_j\|_2, \tag{1.47}$$

for a constant velocity c.

Since the time of the first chirp event is unknown but the period is known, then we can form a measurement matrix z_{ij},

$$\begin{aligned} c(t_{ij} - T_j) &= \|\mathbf{r}_i - \mathbf{s}_j\|_2, && (1.48)\\ \Rightarrow c(t_{ij} - k_1 j - k_0) &= \|\mathbf{r}_i - \mathbf{s}_j\|_2, && (1.49)\\ \Rightarrow c(t_{ij} - k_1 j) &= \|\mathbf{r}_i - \mathbf{s}_j\|_2 + ck_0, && (1.50)\\ \Rightarrow z_{ij} &= \|\mathbf{r}_i - \mathbf{s}_j\|_2 + o. && (1.51) \end{aligned}$$

We can then assume in realistic environments, the errors in the measurements are $\epsilon_{ij} \in N(0, \sigma)$, and that not every measurement will be detected, hence missing information, then we can also define a indexing matrix W with dimensions $m \times n$ whose elements are

$$W_{i,j} = \begin{cases} 1 & \text{if } z_{i,j} \in \mathbb{R}, \\ 0 & \text{otherwise.} \end{cases} \tag{1.52}$$

The problem can be formulated as follows,

$$\tilde{z}_{ij} = \|\mathbf{r}_i - \mathbf{s}_j\|_2 + o + \epsilon_{ij}, \tag{1.53}$$

which in turn can formulate the optimisation problem for parameters $\theta_1 = \{\mathbf{r}, \mathbf{s}, o\}$,

$$\min_{\theta_1} f(\theta_1) = \sum_{(i,j)\in W_{\text{in}}} (z_{ij} - (\|\mathbf{r}_i - \mathbf{s}_j\|_2 + o))^2. \tag{1.54}$$

In a similar spirit to the TOA problem, a compaction matrix $\mathbf{M} = \mathbf{U}^T\mathbf{V}$ can also be formed whose elements are

$$M_{ij} = (z_{ij} - o)^2 - a_i - b_j = u_{i,:}^T v_{:,j}. \tag{1.55}$$

This compaction matrix has a fixed rank of 3 and a relaxed optimization problem can be formed with for parameters $\theta_2 = \{\mathbf{U}, \mathbf{V}, \mathbf{a}, \mathbf{b}, o\}$,

$$\min_{\theta_2} f(\theta_2) = \sum_{(i,j)\in W_{\text{in}}} \left(z_{ij} - (\sqrt{u_i^T v_j + a_i + b_j} + o) \right)^2. \tag{1.56}$$

As part of the research for this problem, a minimal solver was found for $m = n = 5$. This solver gives four solutions for the offset o. A RANSAC method was then created using this minimal solver. Five receivers and five senders were randomly selected, then the four solutions for the offset o is calculated. For each o, the parameters θ_2 were calculated and used to test of inliers in the remaining senders according to equation (2.28). The solution for o, is then used to find more inliers in a RANSAC scheme similar in manner to the TOA RANSAC scheme. Once an inlier set is found, then optimisation using equation (1.54) is performed to refine the parameters of receiver positions $\mathbf{r}$ and sender positions $\mathbf{s}$ in the presence of a unknown constant offset o and to refine the inlier indexing matrix W.

8 Topics for Future Research

The papers studied in this thesis, focus heavily on calibration of receiver positions and sender positions in TOA and related problems. Research efforts were in general, solving these highly non-linear problems when the measurements suffered from errors and missing information. We have also investigated commercially available technologies to simultaneously calibrate receiver positions and sender positions and look into addressing large volume of crowdsourced data. Although these papers have shown ways to overcome these problem, there are always improvements that can be made and other problems arise.

One area in particular I could imagine future research would be a better understanding of IEEE 802.11mc measurements. In our models we assume a Gaussian distribution in the errors. Looking at Figure 1.7, we can see this assumption is not quite true. The measurements seem to show a second peak at a distance further away from the main distribution.

Unfortunately exactly how WiFi modules choose the radio signals in which they use for timing, are known only by the manufactures and not available to the public. This leads me to believe that these measurements may be first order reflected signals. This would imply that classifying these measurements as outliers may not be the correct thing to do, but instead should be classified as reflected signals. These reflected models can have an alternative optimisation function and used to help achieved better estimations. Further to this, it may be possible to estimate room geometry, which would be an interesting area of research. Existing Wifi methods such as fingerprinting do inadvertently achieve a room estimation. Due to the gridded nature of fingerprinting, creating a fingerprint at a location of wall is physically tricky. Other RSSI methods for positioning are usually unreliable due the errors in the measurements. Other radio technologies, such as UWB, are not good for reflections when using Two Way Timing, this would put IEEE 802.11mc at a convenient advantage.

Another area of research would be handling big data. Although a method has be proposed in one of the papers in this thesis, it is by no means the only way. The method proposed still must perform large optimisation steps, but it too has its limits in number of parameters. Research into qualitatively quantifying how good a self calibration estimation without comparing to other maps, would be a step in the right direction. By having this step, it would be much easier to ignore some datasets and find a advantageous selection of maps to create a database of WiFi router locations. At the moment research into large scale self-calibration is very sparse, but hopefully in the near future, this will not be the case. As part of the IEEE 802.11mc protocol, WiFi routers will be able to broadcast there location coordinates to help with positioning, but this is currently changed manually in the routers.

9 Overview of Papers

9.1 Paper I

In the first paper, a framework to robustly solve the TOA self calibration problem with missing information and the presence of outliers in the given data. We proposed a novel hypothesis and test framework that efficiently finds initial estimates of the unknown parameters and combine such methods with optimization techniques to obtain accurate and robust systems. The proposed system was then evaluated using Wi-Fi round-trip time measurements to give a realistic example of indoor localization. This paper was a showcase of how 802.11mc could be used. At the time of this paper, 802.11mc was not commercially released.

Author Contributions:

KÅ and MO conceived and planned the study. KÅ, MO and KB contributed equally to creating the algorithms and experimentation was performed by KB and KÅ equally. The paper was jointly written by KÅ, MO and KB

9.2 Paper II

The second paper is somewhat similar to the first paper. The framework to robustly solve the TOA self calibration problem with missing information and the presence of outliers was refined and more rigorously tested for it's validity. The proposed systems are evaluated against current state-of-the-art methods on a large set of benchmark tests. This is evaluated further on Wi-Fi round-trip time and ultra-wideband measurements to give a realistic examples of self calibration for indoor localization.

Author Contributions:

KÅ and MO conceived and planned the study. KÅ, MO and KB contributed equally to creating the algorithms and experimentation was performed by MO, KB and KÅ equally. The paper was jointly written by KÅ, MO and KB.

9.3 Paper III

In the third paper we study the problem of estimating receiver and sender positions from time-difference-of-arrival measurements, assuming an unknown constant time-difference-of-arrival offset. This problem arises for example with repetitive sound events or a fixed offset on an access point. In this paper it is shown that there are three minimal cases to the problem. One of these (the five receiver, five sender problem) is of particular importance. A fast solver (with run-time under 4 μs) is given. We show how this solver can be used in robust estimation algorithms, based on RANSAC, for obtaining an initial estimate followed by local optimization using a robust error norm. The system is verified on both real sound recordings and synthetic data.

Author Contributions:

KÅ, VL and MO conceived and planned the study. KÅ, VL,MO, GF and KB contributed to creating the algorithms and experimentation was evaluated by MO, VL, GF, KB and KÅ. The sound recordings were performed by TB and HRG in Bulgaria and the sound recordings in Lund were performed by GF, KB and KÅ. The paper was jointly written by all parties.

9.4 Paper IV

In the forth papers a method for performing low rank matrix factorization was proposed. Low-rank matrix factorization is an essential problem in computer vision as described in the introduction. The Low-rank matrix factorizations focused in this paper involve matrices with a specific structure, with large amounts of missing information. Then a RANSAC method was proposed to try multiple iterations of finding usable subsets and expanding the subset. This gives a robust and fast system, with state-of-the-art performance. The system was verified on well known affine structure-from-motion benchmark datasets.

Author Contributions:

MO conceived and planned the study. MO contributed most to creating the algorithms and experimentation was performed by MO, KB and KÅ. The paper was mostly written by MO but contributions were made by KÅ and KB.

9.5 Paper v

This paper follows on nicely from Paper iv. This paper uses many of the methods from the previous papers but to make a real-time implementation. One of the largest hurdles to overcome is the speed of the optimisation, due to the large number of parameters. Too much data, and the optimisation takes too long, too few data and the robust method will not produce useable receiver positions. Hence a middle ground must be found. Its was therefore proposed that batches of data was collected and then matched together. Thankfully for TOA problems using radio, then number of receivers are a lot fewer than the number of sender positions, also the receivers do not move. Due to this reason, the distance matrix is of a fixed structure, which can be taken advanced of. The batches can therefore be 'glued' together using a linear transform, this is then fast and still robust. This then unlocks the potential of having real time calibration. This is then validated using Ultra-Wideband Time-of-Arrival data gathered by a Bitcraze Crazyflie quadcopter in 2D motion, 3D motion and full flight.

Author Contributions:

KB conceived and planned the study following ideas from MO and KÅ. KÅ and MO helped develop the idea. All of the algorithms created, experimentation and evaluation was performed by KB. The paper was mostly written by KB with contributions from KÅ and MO.

9.6 Paper vi

In this paper we take a glimpse at large scale TOA self calibration. With indoor localisation being demanded more, with Internet-Of-Things devices (IoT), a good calibration of receiver positions becomes an infeasible task. The shear number of receivers, let alone good calibrations would be an impossible job when you think about the number of WiFi routers there are in a single office building. With hardware advancements making Ultra-Wideband devices more accurate and low powered, this unlocks the potential of having such devices in commonplace around factories and homes, enabling an alternative method of navigation indoors. Therefore a push for crowd-sourced data acquisition would be necessary, to calibrate such a large number of receivers. In this paper, we presented a method to fuse radio SLAM (also known as Time-Of- Arrival self-calibration) maps together in a linear way. We introduce an automatic scheme to determine which of the maps are best to use, to further improve the receiver calibration. Additionally, when a map is fused in a linear way, it is a very computationally cheap process and produces a reasonable map which is required for crowd-sourced data acquisition.

Author Contributions:

KB conceived and planned the study following ideas from MO and KÅ. KÅ and MO helped develop the idea. All of the algorithms created, experimentation and evaluation was performed by KB. The paper was mostly written by KB with contributions from KÅ and MO.

9.7 Paper VII

In this paper a phase-based positioning framework using a massive MIMO system was presented. The massive MIMO system could track Multipath Components, which can then be used as TOA measurements. The main issue in solving this TOA self calibration problem is dimensionality of the movement of the sender positions. This system is more aligned to a far field approach. However, due to the possibility of multipath components being tracked, the reflected signal give rise to virtual receiver positions, hence improving the dimensionality. The positioning result shows that the mean deviation of the estimated user equipment trajectory from the ground truth is $13cm$, hence the proposed framework promises a high-resolution radio-based positioning solution for current and next generation cellular systems.

Author Contributions:

XL, CG and FT conceived and planned the study. The physical experimentation was performed by XL, CG and FT. All evaluation for solely the Localisation algorithms were conducted by KB, KÅ and MO jointly. All other evaluation was performed by XL, CG and FT. The paper was jointly written by all parties.

Bibliography

[1] H. Stewénius, *Gröbner Basis Methods for Minimal Problems in Computer Vision*. PhD thesis, Lund University, APR 2005.

[2] Y. Kuang, S. Burgess, A. Torstensson, and K. Åström, "A complete characterization and solution to the microphone position self-calibration problem," in *The 38th International Conference on Acoustics, Speech, and Signal Processing*, 2013.

[3] S. Burgess, K. Åström, M. Högström, and B. Lindquist, "Smartphone positioning in multi-floor environments without calibration or added infrastructure," in *2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pp. 1–8, 2016.

[4] G. Young and A. Householder, "Discussion of a set of points in terms of their mutual distances," *Psychometrika*, vol. 3, no. 1, pp. 19–22, 1938.

[5] E. Kaplan and C. Hegarty, *Understanding GPS: principles and applications*. Artech house, 2005.

[6] F. van Diggelen and P. Enge, "The world's first gps mooc and worldwide laboratory using smartphones," in *Proceedings of the 28th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2015)*, pp. 361–369, ION, 2015.

[7] M. Hata, "Empirical formula for propagation loss in land mobile radio services," *IEEE Transactions on Vehicular Technology*, vol. 29, no. 3, pp. 317–325, 1980.

[8] A. A. Adebomehin and S. D. Walker, "Enhanced ultrawideband methods for 5g los sufficient positioning and mitigation," in *2016 IEEE 17th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, pp. 1–4, June 2016.

[9] A. I. Baba, "Calibrating time of flight in two way ranging," in *2011 Seventh International Conference on Mobile Ad-hoc and Sensor Networks*, pp. 393–397, 2011.

[10] "Ieee standard for information technology—telecommunications and information exchange between systems local and metropolitan area networks—specific requirements - part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications," *IEEE Std 802.11-2016 (Revision of IEEE Std 802.11-2012)*, pp. 1–3534, Dec 2016.

[11] A. Shahmansoori, G. E. Garcia, G. Destino, G. Seco-Granados, and H. Wymeersch, "5g position and orientation estimation through millimeter wave mimo," in *2015 IEEE Globecom Workshops (GC Wkshps)*, pp. 1–6, 2015.

[12] E. Björnson, L. Sanguinetti, H. Wymeersch, J. Hoydis, and T. L. Marzetta, "Massive mimo is a reality—what is next?: Five promising research directions for antenna arrays," *Digital Signal Processing*, vol. 94, pp. 3 – 20, 2019. Special Issue on Source Localization in Massive MIMO.

[13] D. Grayson and M. Stillman, "Macaulay 2." Available at http://www.math.uiuc.edu/Macaulay2/, 1993-2002. An open source computer algebra software.

[14] D. Cox, J. Little, and D. O'Shea, *Ideals, Varieties, and Algorithms*. Springer, 2007.

[15] V. Larsson, *Computational Methods for Computer Vision: Minimal Solvers and Convex Relaxations*. PhD thesis, Lund University, 2018.

[16] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, p. 381–395, June 1981.

[17] O. Enqvist, K. Josephson, and F. Kahl, "Optimal correspondences from pairwise constraints," in *Proc. 12th Int. Conf. on Computer Vision, Kyoto, Japan*, 2009.

[18] O. Enqvist, E. Ask, F. Kahl, and K. Åström, "Robust fitting for multiple view geometry," in *European Conference on Computer Vision*, pp. 738–751, Springer, 2012.

[19] C. Tomasi and T. Kanade, "Shape and motion from image streams under orthography: a factorization method," *International journal of computer vision*, vol. 9, no. 2, pp. 137–154, 1992.

[20] C. Eckart and G. Young, "The approximation of one matrix by another of lower rank," *Psychometrika*, vol. 1, no. 3, pp. 211–218, 1936.

[21] K. Levenberg, "A method for the solution of certain non-linear problems in least squares," *Quarterly of applied mathematics*, vol. 2, no. 2, pp. 164–168, 1944.

[22] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *Journal of the society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963.

[23] Y. Kuang and K. Åström, "Stratified sensor network self-calibration from tdoa measurements," in *21st European Signal Processing Conference (EUSIPCO 2013)*, pp. 1–5, 2013.

Chapter 2

Scientific Publications

Paper 1

Robust Time-of-Arrival Self Calibration and Indoor Localization using Wi-Fi Round-Trip Time Measurements

Kenneth Batstone, Magnus Oskarsson and Kalle Åström
Centre of Mathematical Sciences, Lund University, Lund, Sweden

Abstract: The problem of estimating receiver-sender node positions from measured receiver-sender distances is a key issue in different applications such as microphone array calibration, radio antenna array calibration, mapping and positioning using UWB and mapping and positioning using round-trip-time measurements between mobile phones and Wi-Fi-units. Thanks to recent research in this area we have an increased understanding of the geometry of this problem. In this paper, we study the problem of missing information and the presence of outliers in the given data. We propose a novel hypothesis and test framework that efficiently finds initial estimates of the unknown parameters and combine such methods with optimization techniques to obtain accurate and robust systems. The proposed systems are evaluated using Wi-Fi round-trip time measurements to give a realistic example of indoor localization. The resulting map of the anchor points is validated against ground truth measurements with promising results.

1 Introduction

In this paper we present new research on robust methods for time-of-arrival (TOA) self-calibration problem with missing data and outliers. This is then applied to Wi-Fi round-trip time (RTT) indoor localization measurements to provide a realistic example. These measurements are based on the recent IEEE standard 802.11mc (Fine Time Measurement) to measure round-trip time. The TOA self-calibration problem is the problem of determining the positions of a number of receivers and transmitters given receiver-transmitter distances. In this problem, there is no assumption that there exists a subset of sensors (anchors) whose locations are known. Hence it is closely related to the *anchor-free* sensor network localization problem [1] but differs, since the transmitters or receivers are independent of each other. Our problem is bipartite, unlike [1], where the network structure is a general graph. The TOA problem also has certain similarities with the problem of determining a set of points given all inter-point distances, which is usually solved using multi-dimensional scaling [2]. Such problems are of general interest in visualization and analysis of large datasets (e.g. DNA data), in machine learning and for many geometric problems. The TOA self-calibration problem is important for node calibration problems for a variety of different media, e.g. (i) microphone arrays (given recordings of sounds emitted at unknown locations, to microphones at unknown positions, determine both sound emission positions and microphone locations), similarly (ii) ultra-sound, (iii) radio (Ultra Wide Band), (iv) Wi-Fi (Using signal strength) and (v) Wi-Fi (Using RTT distance measurments), which will be used in this paper.

Anchor-free sensor network calibration with time-of-arrival measurements has been investigated in a number of studies. Graph rigidity was explored in [1] to find a fold-free graph embedding. The solution was then refined using mass-spring based optimization. In [3], a semi-definite programming formulation and solution was proposed for TOA measurements, with or without anchors. Both of these methods are general for any solvable[1] network structure, but do not solve the minimal rigid graphs. Such minimal rigid graphs were analyzed and solved in [5]. Another line of work has focused on sensor networks with bipartite structure, that appear in various applications mentioned in the previous paragraph. For this special type of bipartite network structure, one also aims to identify and solve the minimal problem, i.e. minimal number of receivers and transmitters required for the problem to be well-defined (or solvable).These minimal problems can serve as initial solutions in the robust estimation framework. Note that for this problem, the roles of receivers and transmitters are equivalent. Therefore, when discussing minimal cases, the number of sensors required for receivers and transmitters are interchangeable. The minimal cases were studied in [6], where solutions to the minimal case of 3 transmitters and 3 receivers in the plane are given. The minimal problems for the 3D case are given in [7]. The minimal num-

[1] when the graph is globally rigid cf. [4]

ber of receivers and senders are $(4, 6)$, $(5, 5)$ and $(6, 4)$ respectively. There are in general 38, 42 and 38 solutions respectively for the three types of problems. In the paper, there are also practical algorithms for the solution to the problem. Previous to this the state of the art in terms of practical algorithms were that of [8], where a non-minimal linear solution to 3D TOA self-calibration problem is derived for the combination of $(4, 10)$ receivers and senders. Other relevant results on determining the positions from distances are [9], where synchronization between the receivers and senders are not assumed, [10] where pose for dual receiver rigs are considered, [11, 12], where the TOA self-calibration problem when say the receivers span a linear space of lower dimension than that of the senders, and [13], where minimal TOA self-calibration problems are studied and solved for the case of additional distance measurements among e.g. the receivers.

In [14, 15] a solution is given to the TOA self-calibration problem, if one may additionally assume that one of the receivers coincides with the position of one of the transmitters. The minimal cases for far field approximation was first studied in [16] and later refined in [17]. In far field approximation, the distances between the transmitters and receivers are assumed to be considerably larger than those between receivers. The solutions based on the far field approximation can be utilized to initialize the original TOA problem. Studying these minimal cases is of both theoretical importance and essential to develop fast and stable algorithms suitable for robust estimation methods like RANSAC [18] in the presence of outliers in the measurements.

As will be shown in the following sections, one part of our system exploits that the so-called *compaction matrix* should have a certain rank. Low rank matrix factorization has a long standing history. Truncating the singular value decomposition of the measurement matrix has been shown to give the optimal solution under the l^2-norm when for complete data, see [19]. The work in [20] was the first to consider missing data. Robustness to outliers has been considered in [21, 22, 23, 24]. Most methods mentioned above are based on alternating optimization and are prone to get trapped in local minima. Recently, several works [25, 26, 27] re-formulate the problem to minimize the convex surrogate of the rank function, that is, the nuclear norm. For applications with a given fixed rank, the nuclear norm based methods usually perform inferior to the bilinear formulation-based methods [28]. A few recent works [29, 30] also explore the idea to divide the whole matrix into overlapping sub-blocks and combine the sub-block solutions. Minimal cases for low rank matrix factorization, for missing data, were investigated in [31].

The novel concept of this paper, is that we can use the hypothesis and test framework to find a robust initial estimates. By using the proposed technique, it has the advantage of avoiding using missing data and robustly identifying outlier data. It is also a very fast algorithm, which takes 0.588375 seconds for the 2D planar case on a single core of a Intel Core 2 Duo E7500 2.93 GHz processor and can easily be made faster by parallel computing.

2 Basic Geometry

We will now describe the basic underlying geometry of our problem. Let $\mathbf{r}_i$, $i = 1, \ldots, m$ and $\mathbf{s}_j$, $j = 1, \ldots, n$ be the spatial coordinates of m receivers (e.g. microphones) and n transmitters (e.g. sound events), respectively. For measured time of arrival t_{ij} from transmitter $\mathbf{r}_i$ and receiver $\mathbf{s}_j$, we have $vt_{ij} = \|\mathbf{r}_i - \mathbf{s}_j\|_2$ where v is the speed of measured signals and $\|.\|_2$ is the l^2-norm. The speed v is assumed to be known and constant. We further assume that we, at each receiver can distinguish which transmitter j each event is originating from. This can be done e.g. if the signals are temporally separated or using different frequencies. We will in the following work with the distance measurements $d_{ij} = vt_{ij}$. It is quite common that such data contains both missing data (not every sound event is detected at every microphone) and outliers (e.g. due to errors in the matching process). The TOA calibration problem can then be defined as follows,

Problem 1. *(Time-of-Arrival Self-Calibration) Given absolute distance measurements*

$$d_{ij} = \|\mathbf{r}_i - \mathbf{s}_j\|_2 + \epsilon_{ij}, \tag{2.1}$$

for a subset $W \subset I$ of all the receiver-transmitter index pairs $I = \{(i,j)|i = 1, \ldots m, j = 1, \ldots, n\}$ determine receiver positions $\mathbf{r}_i$, $i = 1, \ldots, m$ and transmitter positions $\mathbf{s}_j$, $j = 1, \ldots, n$. Here the errors ϵ_{ij} are assumed to be either **inliers***, in which case the errors are small ($\epsilon_{ij} \in N(0, \sigma)$) or* **outliers***, in which case the measurements are way off.*

Here we will use the set W_i for the indices (i,j) corresponding to the inlier measurements and W_o for the indices corresponding to the outlier set.

We will now show how the TOA calibration problem is solved generally. From many types of media, a transmitter-receiver distance will be acquired, d_{ij}. Since this can be assumed to be real and positive, it can be squared as follows,

$$d_{ij}^2 = (\mathbf{r}_i - \mathbf{s}_j)^T(\mathbf{r}_i - \mathbf{s}_j) = \mathbf{r}_i^T\mathbf{r}_i + \mathbf{s}_j^T\mathbf{s}_j - 2\mathbf{r}_i^T\mathbf{s}_j. \tag{2.2}$$

The problem is then reformed according to the following invertible linear combinations of d_{ij}^2:

$$\mathbf{B} = \begin{pmatrix} d_{11}^2 & d_{12}^2 - d_{11}^2 & \ldots & d_{1n}^2 - d_{11}^2 \\ d_{21}^2 - d_{11}^2 & & & \\ \ldots & & \hat{\mathbf{B}} & \\ d_{m2}^2 - d_{11}^2 & & & \end{pmatrix}, \tag{2.3}$$

where the *compaction matrix* $\hat{\mathbf{B}}$ is an $(m-1) \times (n-1)$ matrix with entries as $\hat{B}_{ij} = \frac{d_{ij}^2 - d_{i1}^2 - d_{1j}^2 + d_{11}^2}{-2}$, with $i = 2, \ldots m$ and $j = 2, \ldots, n$. The other elements in $\mathbf{B}$ are used as constraints for the solution.

The factorization can then be interpreted as follows. Let $\mathbf{R}_i = \left[(\mathbf{r}_i - \mathbf{r}_1)\right]$ and $\mathbf{S}_j = \left[(\mathbf{s}_j - \mathbf{s}_1)\right]$. Here $\hat{\mathbf{B}} = \mathbf{R}^T\mathbf{S}$ with $\mathbf{R}_i$ as columns of $\mathbf{R}$ and $\mathbf{S}_j$ as columns of $\mathbf{S}$. Since we assume that $\mathbf{R}$ and $\mathbf{S}$ are in a 3D affine space, the matrix $\hat{\mathbf{B}}$ has rank 3 at most. This also implies that in order to solve the problem, it is required that $m \geq 4$ and $n \geq 4$.

By factorizing $\hat{\mathbf{B}}$, we can compute the vectors to all receivers and transmitters from unknown initial/reference positions ($\mathbf{r}_1$ and $\mathbf{s}_1$).

By fixing $\mathbf{r}_1$ at the origin and $\mathbf{s}_1$ as a vector from the origin, in terms of an affine transformation matrix $\mathbf{L}$ and vector $\mathbf{b}$, the problem is reformulated as follows,

$$\begin{aligned} &\mathbf{r}_1 = 0,\ \mathbf{s}_1 = \mathbf{L}\mathbf{b}, \mathbf{r}_i = \mathbf{L}^{-T}\tilde{\mathbf{R}}_i,\ i = 2 \ldots m, \\ &\mathbf{s}_j = \mathbf{L}(\tilde{\mathbf{S}}_j + \mathbf{b}),\ j = 2 \ldots n, \end{aligned} \tag{2.4}$$

where $\tilde{\mathbf{R}} = \mathbf{L}^T\mathbf{R}$, $\tilde{\mathbf{S}} = \mathbf{L}^{-1}\mathbf{S}$, and hence $\hat{\mathbf{B}} = \tilde{\mathbf{R}}^T\mathbf{L}^{-1}\mathbf{L}\tilde{\mathbf{S}} = \mathbf{R}^T\mathbf{S}$.

Using this parametrization, the equations from matrix $\mathbf{B}$, (2.46) become

$$\begin{aligned} d_{11}^2 &= (\mathbf{r}_1 - \mathbf{s}_1)^T(\mathbf{r}_1 - \mathbf{s}_1) = \mathbf{s}_1^T\mathbf{s}_1 = \mathbf{b}^T\mathbf{L}^T\mathbf{L}\mathbf{b} \\ &= \mathbf{b}^T\mathbf{H}^{-1}\mathbf{b}, \qquad (2.5) \\ d_{1j}^2 - d_{11}^2 &= \mathbf{s}_j^T\mathbf{s}_j - \mathbf{s}_1^T\mathbf{s}_1 = \tilde{\mathbf{S}}_j^T\mathbf{L}^T\mathbf{L}\tilde{\mathbf{S}}_j + 2\mathbf{b}^T\mathbf{L}^T\mathbf{L}\tilde{\mathbf{S}}_j \\ &= \tilde{\mathbf{S}}_j^T\mathbf{H}^{-1}\tilde{\mathbf{S}}_j + 2\mathbf{b}^T\mathbf{H}^{-1}\tilde{\mathbf{S}}_j, \qquad (2.6) \\ d_{i1}^2 - d_{11}^2 &= \mathbf{r}_i^T\mathbf{r}_i - 2\mathbf{r}_i^T\mathbf{s}_1 = \tilde{\mathbf{R}}_i^T(\mathbf{L}^T\mathbf{L})^{-1}\tilde{\mathbf{R}}_i - 2\mathbf{b}^T\tilde{\mathbf{R}}_i \\ &= \tilde{\mathbf{R}}_i^T\mathbf{H}\tilde{\mathbf{R}}_i - 2\mathbf{b}^T\tilde{\mathbf{R}}_i, \qquad (2.7) \end{aligned}$$

where the symmetric matrix $\mathbf{H} = (\mathbf{L}^T\mathbf{L})^{-1}$. With this parameterization, there are in total 9 unknowns (6 and 3 unknowns for $\mathbf{H}$ and $\mathbf{b}$, respectively), and hence a solution can be found. Since this solution has its own coordinate system, with prior knowledge this can be transformed back to the original coordinate system. For a 2D problem, there are a total of 5 unknowns (3 and 2 unknowns for $\mathbf{H}$ and $\mathbf{b}$, respectively).

3 Non-Linear Optimization Approaches

In the development of the different systems for robust estimation, we use several different local optimization techniques. In particular we use methods for local optimization of the type

$$\min_{\mathbf{r},\mathbf{s}} \sum_{(i,j)\in\tilde{W}} f(d_{ij} - ||\mathbf{r}_i - \mathbf{s}_j||_2), \tag{2.8}$$

where $f(r)$ is chosen to be (i) $f(r) = r^2$ (l^2-norm), (ii) $f(r) = |r|$ (l^1-norm) or (iii) $f(r) = \min(r^2, T)$ (truncated l^2-norm) and $\tilde{W}$ is a subset of the data measurements. If the subset $\tilde{W}$ contains no outliers and if the starting point is good, then the l^2-norm can give good estimates. Optimizing using the l^1-norm is less sensitive to the subset $\tilde{W}$ containing outliers, but still require a reasonably good starting point to converge to a good solution. Local optimization of the truncated l^2-norm is even more sensitive to having a good starting point. Nevertheless, these local optimization techniques are important components for designing robust systems.

4 Obtaining Initial Estimates

Finding the optimal solution to problem 5, in the presence of outliers and missing data is a highly non-convex problem. We are thus dependent on finding good initial starting solutions, for the optimization methods from the previous section to work. We will, in this section, describe the different initialization methods that we have explored. In the next section we will describe our main contribution to the initialization problem.

Arguably, the most straight-forward way to initialize a solution, is to simply randomly place all receivers and senders within some space. This usually gives poor initial estimates, and the local optimization will be prone to get stuck in local minima. A slight improvement to this idea, is to use multiple restarts and optimize from each initial position, and then in the end choose the best solution.

Another way of initializing, that we have explored, is using the rank constraint on the compaction matrix. Here one can use many existing methods for doing the low rank matrix factorization. One important draw-back of these methods, is that we need to have at least one row and one column of the data matrix completely known, and without outliers. The last criteria is of course hard to to check. If all data is known, the optimal low rank factorization is given by singular value decomposition (SVD) of the data matrix. A heuristic for handling missing data, is simply to fill in the missing data with some random values that follow the statistics of the other known measurements. One can then use SVD to obtain an initial estimate. This can be used directly to find the solution to the original problem as described in section 2. Alternatively, the initial low rank matrix factorization can be refined using the Wiberg algorithm, [20].

5 Random Sampling Paradigm

The RANSAC or hypothesize and test paradigm, has proven to be useful in situations where there are outliers in the data, [18]. In this paradigm, a subset of the data is used to estimate

the unknown parameters. The remainder of the data is then used to verify or falsify the parameters. This is typically repeated a fixed number of iterations. The parameters that give the largest number of inliers are then usually used as an initial estimate for the subsequent non-linear optimization of the parameters.

For Problem 5, there are several ways one could implement the hypothesis and test paradigm. One idea would be to use efficient algorithms for determining receiver and sender positions from minimal data, [7]. The minimal number of receivers and senders are $(4, 6)$, $(5, 5)$ and $(6, 4)$ respectively. There are in general 38, 42 and 38 solutions respectively for the three types of 3D problems. For the 2D problem in this experiment, the minimal number of receivers and senders are $(3, 3)$. One way of using the $(4, 6)$ solver would be to randomly select 4 receivers, then randomly select 6 senders such that there are no missing data for this particular receiver-sender configuration. From each of the solutions to the minimal problems one obtains a hypothesis of the positions of the 4 receivers. The test step would then consist of checking for how many of the remaining senders it is possible to find a position that fits all of the 4 distance measurements.

Although the $(4, 6)$-solver and the test is relatively fast, we propose an alternative to this approach. The main idea is to find a fast way to hypothesize and test. We will use the rank constraints of the compaction matrix to do this. Our method is described in Algorithm 2.

Algorithm 1 Our RANSAC initialization scheme

1: Select 5 receivers randomly
2: Find all senders, for which there are no missing data to the 5 receivers
3: If there are at least five such senders, select 4 of these senders randomly
4: **Hypothesize:** Use the 5×4 matrix $\mathbf{F}$, with elements $\mathbf{F}_{ij} = d_{ij}^2$. Calculate the compaction matrix $\hat{\mathbf{B}}$ for $\mathbf{F}$ as shown in equation (2.46), which is a 4×3 matrix. Calculate a unit vector $\mathbf{v}$ which lies in the left null space of $\hat{\mathbf{B}}$.
5: **Test:** Assuming that a column $\mathbf{f}$, whose elements are $\mathbf{f}_i = d_{ij_{test}}^2$ contain no outliers, then the vector $\hat{\mathbf{b}}$ is the compaction matrix of $[\mathbf{F}_1, \mathbf{f}]$ where $\mathbf{F}_1$ is the first column of $\mathbf{F}$. This then should have $\mathbf{v} \cdot \hat{\mathbf{b}} = 0$. Assuming low noise it is reasonable to declare it an inlier if $|\mathbf{v} \cdot \hat{\mathbf{b}}| < T$, where T is a threshold that depends on the noise level σ and the data d. Repeat this test for all the other columns.
6: Repeat steps 1-5, K times and keep track of the hypothesis that gave the largest number of inlier columns.

6 Experimental Setup

Indoor localization is a currently a key issue, from needing to know the location of objects using Ultra-Wide Band beacons to finding the location of mobile phones when a GPS

signal cannot be acquired, this is prevalent indoors and in built up areas such as New York, which is known as the "urban canyon" problem.

To test our system we decided to use 5 Nexus 6 mobile phones, 4 as Wi-Fi anchors and 1 as a moving source to give a realistic scenario of a wireless network. These phones come with a IEEE 802.11.mc Wi-Fi standard, which also produces round-trip time measurements in metres. This Wi-Fi standard is set to be used widely next year. When conducting the experiment we used a Wi-Fi frequency of 5 GHz on channel number 44 with a 80 MHz bandwidth.

The experimental environment that was chosen was a large open space in an office block (Ideon Alfahuset, Lund, Sweden) with dimensions $\sim 12 \times 18$ m. This gave the advantage of having a small amount of reflecting surfaces, the only notable obstruction was a staircase obscuring the direct line of sight between the anchors and the source, please see the schematic of the experiment Figure 2.1.

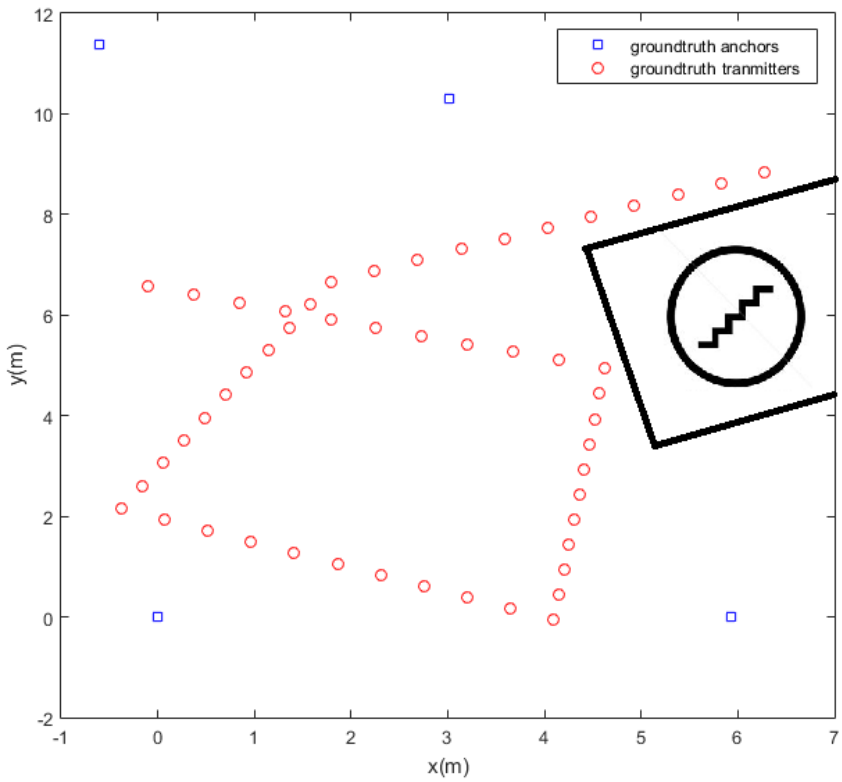


Figure 2.1: A diagram of the experimental setup with the staircase illustrated on the right side of the image.

For the experiment, a ground truth was measured for all the anchor points and the path of the moving source every $0.5 \pm 0.005m$ with a tape measure. The anchors and the moving source were risen from the floor using stands and kept at the same height in order for the experiment to be conducted on a single plane.

The experiment was conducted by walking the predefined path with the moving source on a stand and trying to ensure that the body of the person walking the path was not in the way of the line of sight to any of the anchors. Currently, when the moving source is taking a distance measurements, this can only be done recursively to each of the anchors, one at a time. For this reason, when conducting the experiment, small pauses of ~ 2 *seconds* were taken at each of the ground truth points to ensure that there was an opportunity for distance measurements to be taken to each of the anchors.

In our experiment, our proposed framework does not require prior knowledge about the positions for any of the anchors, positions of any of the source locations and no prior knowledge of the number of anchors and source locations. The only requirement is that the number of anchors and source locations satisfy the requirement of the minimal solver (i.e. (3,3)).

7 Experimental Evaluation

Once the measurements were taken, the data was filtered using a status variable from the 802.11.mc standard in which a good RTT measurement is 0. The other values were replaced with a not a number (NaN) value. This can then be used in our RANSAC initialization scheme, Algorithm 2. For this experiment we used the (3,3) receivers and senders minimal solver. A fixed number of iterations was used; 20 iterations for the initial selection of 3 receivers and senders, then a further 70 iterations to extend the number of columns and another 70 to extend the number of rows. The tolerance was set to $T = 3$ for the initial selection and extension of columns then reduced to $T = 2$ for the extensions of the rows.

Now that the initial values have been estimated, it undergoes l^2 optimization on the inlier set. We also add a smoothness prior in the optimization. This prior is based on minimizing acceleration, according to

$$res_a = \frac{1}{\sigma_a^2} \sum_{j=2}^{n-1} \|\mathbf{s}_{j-1} - 2\mathbf{s}_j + \mathbf{s}_{j+1}\|_2^2, \tag{2.9}$$

where σ_a is a parameter controlling the strength of the smoothness prior. The results of the estimated anchor positions and source positions after the optimization are shown in Figure 2.5.

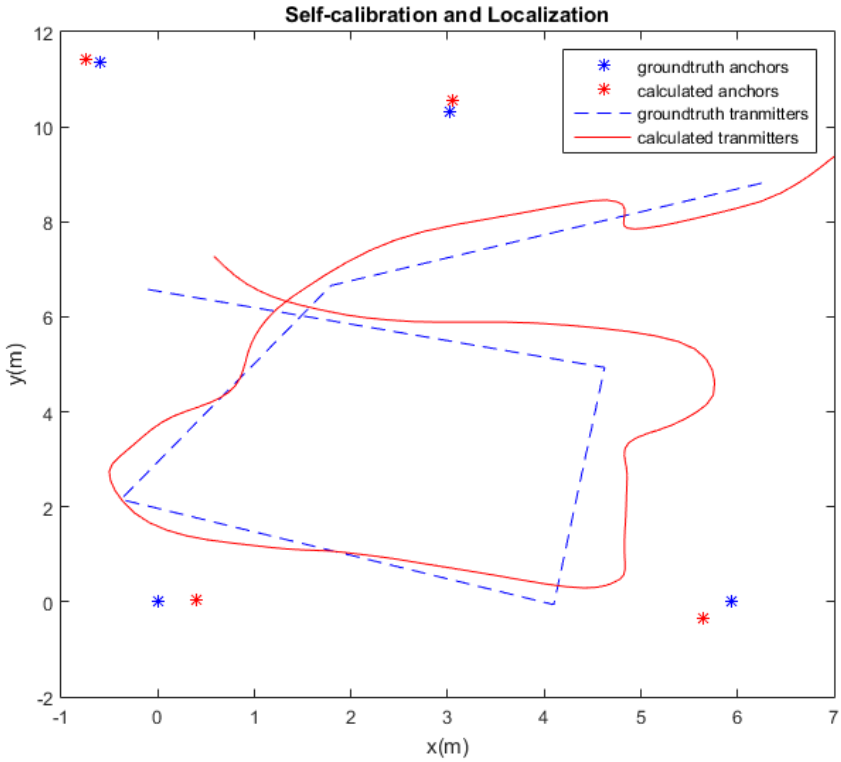


Figure 2.2: This figure illustrates the estimated anchor positions and the source positions. This is overlaid on the ground truth anchor positions and predefined path.

The ground truth anchor positions were $((0,0),(5.932,0),(3.0163,10.3079),(-0.6018,11.3658))$. The estimated anchors were calculated to be $((0.4044,0.0429),(5.6435,-0.3461),(3.0475,10.5527),(-0.7488,11.4242))$. This gives a total Euclidean distance error (i.e. the 2-norm of the difference in positions) of 0.5923. For graphical purposes, a Procrustes fitting was used on the anchor positions to spread the total error over all 4 anchors in stead of just 3 anchors. In the Procrustes fitting only rotation and translation were allowed.

To calculate the errors in the estimated source positions, the Euclidean distance from the estimated source to the nearest location in ground truth set was calculated. The ground truth data was linearly interpolated in order for the Euclidean distance to be as close as possible to the predefined path. The residuals are shown in Figure 2.3.

For just the section of the predefined path that obscured by the staircase. The measurements had 6.52% missing data and 9.7826% of the data was considered to be outliers, according to our algorithm. In comparison to the whole data set we had 1.23% missing data and

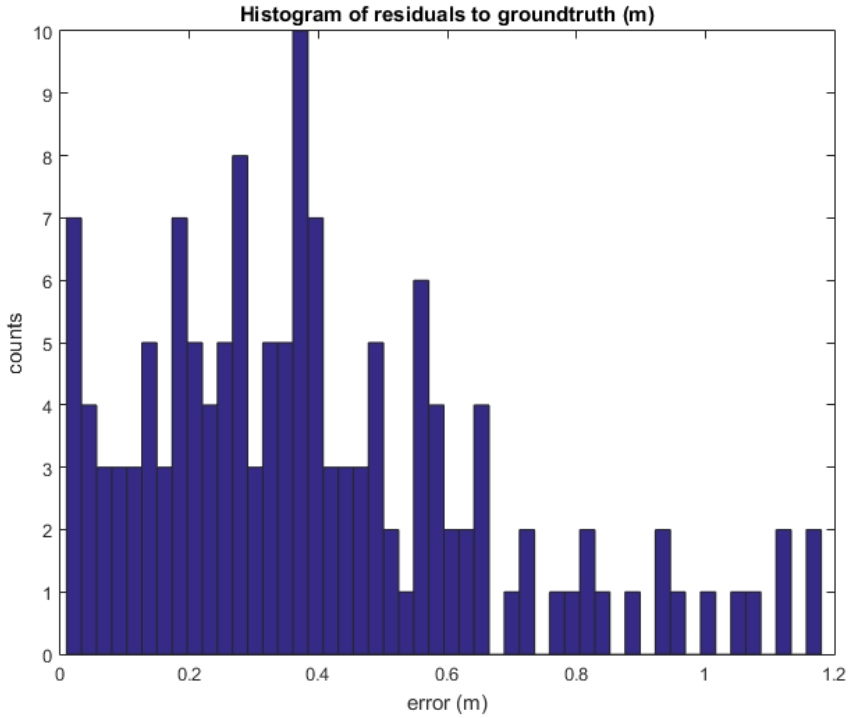


Figure 2.3: The residual error from the estimated source positions and ground truth predefined path.

3.3451% of the data was considered to be outliers.

8 Conclusion

In this paper we have constructed a novel random sampling paradigm system for estimating receiver-sender node positions from measured distance with outliers and missing data. This was verified using Wi-Fi RTT data. Looking at Figure 2.5, we can see that a reasonable estimation was calculated. The general predefined path shape was preserved and the anchor positions are reasonable with a total error of 0.5923 m. From a previous experiment we conducted with 1 anchor and a single source position with direct line of sight, we found that the distance measurement would be a normal distribution with a standard deviation of 1.08 m. If we then compare this value to the total error in the anchor positions and the calculated source positions in Figure 2.3, it can be concluded that our estimations are accurate.

Further to this, the predefined path goes past the set of stairs, this would cause the anchor

to be obscured so the measurements would become either outliers or a missed data point. Despite this, the proposed method robustly and accurately estimated source positions. The experimental results demonstrates the robustness of the proposed method in this Wi-Fi field and how it can be effective in other areas as the errors in the estimations depend predominately on the accuracy of the distance measurements.

Acknowledgment

The authors would like to thank MAPCI and ELLIIT for funding. The authors would also like to thank Björn Lindquist and Rasmus Ljungberg at Combain AB, Lund, Sweden for their expertise and help in conducting the experiments.

Bibliography

[1] N. Priyantha, H. Balakrishnan, E. Demaine, and S. Teller, "Anchor-free distributed localization in sensor networks," in *Proceedings of the 1st international conference on Embedded networked sensor systems*, pp. 340–341, ACM, 2003.

[2] G. Young and A. Householder, "Discussion of a set of points in terms of their mutual distances," *Psychometrika*, vol. 3, no. 1, pp. 19–22, 1938.

[3] P. Biswas, T.-C. Lian, T.-C. Wang, and Y. Ye, "Semidefinite programming based algorithms for sensor network localization," *ACM Trans. Sen. Netw.*, vol. 2, pp. 188–220, May 2006.

[4] T. Eren, O. Goldenberg, W. Whiteley, Y. Yang, A. Morse, B. Anderson, and P. Belhumeur, "Rigidity, computation, and randomization in network localization," in *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, vol. 4, pp. 2673–2684, IEEE, 2004.

[5] M. Oskarsson, K. Åström, and A. Torstensson, "Prime rigid graphs and multidimensional scaling with missing data," in *Pattern Recognition (ICPR), 2014 22nd International Conference on*, pp. 750–755, Aug 2014.

[6] H. Stewénius, *Gröbner Basis Methods for Minimal Problems in Computer Vision*. PhD thesis, Lund University, APR 2005.

[7] Y. Kuang, S. Burgess, A. Torstensson, and K. Åström, "A complete characterization and solution to the microphone position self-calibration problem," in *The 38th International Conference on Acoustics, Speech, and Signal Processing*, 2013.

[8] M. Pollefeys and D. Nister, "Direct computation of sound and microphone locations from time-difference-of-arrival data," in *Proc. of International Conference on Acoustics, Speech and Signal Processing*, 2008.

[9] S. Burgess, Y. Kuang, and K. Åström, "Node localization in unsynchronized time of arrival sensor networks," in *Proceedings of the 21st International Conference on Pattern Recognition*, 2012.

[10] S. "Burgess, Y. Kuang, and K. Åström, "Pose estimation from minimal dual-receiver configurations," in *Proceedings of the 21st International Conference on Pattern Recognition*, 2012.

[11] S. Burgess, Y. Kuang, and K. Åström, "Toa sensor network calibration for receiver and transmitter spaces with difference in dimension," in *Proceeding of the 21st European Signal Processing Conference 2013*, 2013.

[12] S. Burgess, Y. Kuang, and K. Åström, "Toa sensor network self-calibration for receiver and transmitter spaces with difference in dimension," *Signal Processing*, vol. 107, pp. 33–42, 2015.

[13] S. Zhayida, S. Burgess, Y. Kuang, and K. Åström, "Minimal solutions for dual microphone rig self-calibration," in *European Signal Processing Conference (EUSIPCO 2014)*, 2014.

[14] M. Crocco, A. Del Bue, M. Bustreo, and V. Murino, "A closed form solution to the microphone position self-calibration problem," in *37th International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2012), Kyoto, Japan*, March 2012.

[15] M. Crocco, A. Del Bue, and V. Murino, "A bilinear approach to the position self-calibration of multiple sensors," *Trans. Sig. Proc.*, vol. 60, pp. 660–673, feb 2012.

[16] S. Thrun, "Affine structure from sound," in *Proceedings of Conference on Neural Information Processing Systems (NIPS)*, (Cambridge, MA), MIT Press, 2005.

[17] Y. Kuang, E. Ask, S. Burgess, and K. Åström, "Understanding toa and tdoa network calibration using far field approximation as initial estimate," in *ICPRAM*, 2012.

[18] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–95, 1981.

[19] C. Eckart and G. Young, "The approximation of one matrix by another of lower rank," *Psychometrika*, vol. 1, no. 3, pp. 211–218, 1936.

[20] T. Wiberg, "Computation of principal components when data are missing," in *Proc. Second Symp. Computational Statistics*, 1976.

[21] H. Aanaes, R. Fisker, K. Åström, and J. Carstensen, "Robust factorization," *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2002.

[22] A. M. Buchanan and A. W. Fitzgibbon, "Damped newton algorithms for matrix factorization with missing data," in *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 2 - Volume 02*, CVPR '05, (Washington, DC, USA), pp. 316–322, IEEE Computer Society, 2005.

[23] Q. Ke and T. Kanade, "Robust L_1 norm factorization in the presence of outliers and missing data by alternative convex programming," in *Proc. Conf. Computer Vision and Pattern Recognition*, 2005.

[24] A. Eriksson and A. Hengel, "Efficient computation of robust weighted low-rank matrix approximations using the L_1 norm," *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2012.

[25] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?," *Journal of ACM*, 2009.

[26] R. Garg, A. Roussos, and L. Agapito, "Dense variational reconstruction of non-rigid surfaces from monocular video," in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pp. 1272–1279, IEEE, 2013.

[27] C. Olsson and M. Oskarsson, "A convex approach to low rank matrix approximation with missing data," in *Scandinavian Conf. on Image Analysis*, 2011.

[28] R. Cabral, F. D. la Torre, J. P. Costeira, and A. Bernardino, "Unifying nuclear norm and bilinear factorization approaches for low-rank matrix decomposition," in *Proc. Int. Conf. on Computer Vision*, 2013.

[29] L. W. Mackey, M. I. Jordan, and A. Talwalkar, "Divide-and-conquer matrix factorization," in *Advances in Neural Information Processing Systems*, pp. 1134–1142, 2011.

[30] V. Larsson, C. Olsson, E. Bylow, and F. Kahl, "Rank minimization with structured data patterns," in *ECCV*, 2014.

[31] F. Jiang, M. Oskarsson, and K. Åström, "On the minimal problems of low-rank matrix factorization," in *Proc. Conf. Computer Vision and Pattern Recognition*, 2015.

Paper II

Robust Time-of-Arrival Self Calibration with Missing Data and Outliers

Kenneth Batstone, Magnus Oskarsson and Kalle Åström
Centre of Mathematical Sciences, Lund University, Lund, Sweden

Abstract: The problem of estimating receiver-sender node positions from measured receiver-sender distances is a key issue in different applications such as microphone array calibration, radio antenna array calibration, mapping and positioning using ultra-wideband and mapping and positioning using round-trip-time measurements between mobile phones and Wi-Fi-units. Thanks to recent research in this area we have an increased understanding of the geometry of this problem. In this paper, we study the problem of missing information and the presence of outliers in the data. We propose a novel hypothesis and test framework that efficiently finds initial estimates of the unknown parameters and combine such methods with optimization techniques to obtain accurate and robust systems. The proposed systems are evaluated against current state-of-the-art methods on a large set of benchmark tests. This is evaluated further on Wi-Fi round-trip time and ultra-wideband measurements to give a realistic example of self calibration for indoor localization.

1 Introduction

In this paper we present new research on robust methods for time-of-arrival (TOA) self-calibration problem with missing data and outliers. This is then applied to Wi-Fi round-trip time (RTT) and ultra-wideband indoor localization measurements to provide realistic examples. The TOA self-calibration problem is the problem of determining the positions of a number of receivers and transmitters given receiver-transmitter distances. In this problem, there is no assumption that there exists a subset of sensors (anchors) whose locations are known. Hence it is closely related to the *anchor-free* sensor network localization problem [1] but differs, since the transmitters or receivers are independent of each other. Our problem structure corresponds to a bipartite graph, unlike [1] where the network structure is a general graph. The TOA problem also has certain similarities with the problem of determining a set of points given all inter-point distances, which is usually solved using multi-dimensional scaling [2]. Such problems are of general interest in visualization and analysis of large datasets (e.g. DNA data), in machine learning and for many geometric problems. The TOA self-calibration problem is important for node calibration problems for a variety of different media, e.g. (i) microphone arrays (given recordings of sounds emitted at unknown locations, to microphones at unknown positions, determine both sound emission positions and microphone locations), similarly (ii) ultra-sound, (iii) radio (Ultra Wide Band), (iv) Wi-Fi (Using signal strength) and (v) Wi-Fi round-trip time measurements (RTT).

Anchor-free sensor network calibration with time-of-arrival measurements has been investigated in a number of studies. Graph rigidity was explored in [1] to find a fold-free graph embedding. The solution was then refined using mass-spring based optimization. In [3], a semi-definite programming formulation and solution was proposed for TOA measurements, with or without anchors. Both of these methods are general for any solvable network structure. Another line of work has focused on sensor networks with bipartite structure, that appear in various applications mentioned in the previous paragraph. For this special type of bipartite network structure, one also aims to identify and solve the minimal problem, i.e. minimal number of receivers and transmitters required for the problem to be well-defined (or solvable). Note that for this problem, the roles of receivers and transmitters are equivalent. Therefore, when discussing minimal cases, the number of sensors required for receivers and transmitters are interchangeable. The minimal cases were studied in [5], where solutions to the minimal case of 3 transmitters and 3 receivers in the plane are given. The minimal problems for the 3D case are given in [6]. The minimal number of receivers and senders are $(4, 6)$, $(5, 5)$ and $(6, 4)$ respectively. There are in general 38, 42 and 38 solutions respectively for the three types of problems. However, no practical methods for general 3D positions are given. There are a few results on algorithms for actually determining the positions from distances, most notably [7, 8]. In [9], a non-minimal

Thanks to MAPCI and ELLIIT for funding.
when the graph is globally rigid cf. [4]

linear solution to the 3D TOA self-calibration problem is derived for 10 (4) receivers and 4 (10) transmitters. In [10, 11] a solution is given to the TOA self-calibration problem, if one may additionally assume that one of the receivers coincides with the position of one of the transmitters. The minimal cases for far field approximation were first studied in [12] and later refined in [13]. In far field approximation, the distances between the transmitters and receivers are assumed to be considerably larger than those between receivers. The solutions based on the far field approximation can be utilized to initialize the original TOA problem.

Studying these minimal cases is of theoretical importance and further more essential when developing fast and stable algorithms based on robust estimation methods like RANSAC [14], in the presence of outliers in the measurements. As will be shown in the following sections, one important part of our system exploits that the so-called *compaction matrix* should have a certain rank. Low rank matrix factorization has a long standing history. Truncating the singular value decomposition of the measurement matrix has been shown to give the optimal solution under the l^2-norm for complete data, see [15]. The work in [16] was the first to consider missing data. Robustness to outliers has been considered in [17, 18, 19, 20]. Most methods mentioned above are based on alternating optimization and are prone to get trapped in local minima. Recently, several works [21, 22, 23] re-formulate the problem to minimize the convex surrogate of the rank function, that is, the nuclear norm. For applications with a given fixed rank, the nuclear norm based methods usually perform inferior to the bilinear formulation-based methods [24]. A few recent works [25, 26] also explore the idea to divide the whole matrix into overlapping sub-blocks and combine the sub-block solutions. Minimal cases for low rank matrix factorization, for missing data, were investigated in [27].

Indoor localization is a currently a key issue, from needing to know the location of objects using Ultra-Wide Band beacons to finding the location of mobile phones with Wi-Fi when a GPS signal cannot be acquired. This is prevalent indoors and in build up areas such as New York, which is known as the "urban canyon" problem. Methods, like the one proposed, could be useful in solving these real world problems.

2 Basic Geometry

We will now describe the basic underlying geometry of our problem. Let $\mathbf{r}_i$, $i = 1, \ldots, m$ and $\mathbf{s}_j$, $j = 1, \ldots, n$ be the spatial coordinates of m receivers (e.g. microphones) and n transmitters (e.g. sound events), respectively. For measured time of arrival t_{ij} from transmitter $\mathbf{r}_i$ and receiver $\mathbf{s}_j$, we have $vt_{ij} = \|\mathbf{r}_i - \mathbf{s}_j\|_2$ where v is the speed of measured signals and $\|.\|_2$ is the l^2-norm. The speed v is assumed to be known and constant. We further assume that we, at each receiver can distinguish which transmitter j each event is originating from. This can be done e.g. if the signals are temporally separated or using different

frequencies. We will in the following work with the distance measurements $d_{ij} = vt_{ij}$. It is quite common that such data contains both missing data (not every sound event is detected at every microphone) and outliers (e.g. due to errors in the matching process). The TOA calibration problem can then be defined as follows,

Problem 2. *(Time-of-Arrival Self-Calibration) Given absolute distance measurements*

$$d_{ij} = \|\mathbf{r}_i - \mathbf{s}_j\|_2 + \epsilon_{i,j}, \tag{2.10}$$

for a subset $W \subset I$ of all the receiver-transmitter index pairs $I = \{(i,j)|i = 1,\dots m, j = 1,\dots,n\}$ determine receiver positions $\mathbf{r}_i$, $i = 1,\dots,m$ and transmitter positions $\mathbf{s}_j$, $j = 1,\dots,n$. Here the errors $\epsilon_{i,j}$ are assumed to be either **inliers***, in which case the errors are small ($\epsilon_{i,j} \in N(0,\sigma)$) or* **outliers***, in which case the measurements are way off.*

Here we will use the set W_i for the indices (i,j) corresponding to the inlier measurements and W_o for the indices corresponding to the outlier set.

We will now show how the TOA calibration problem is solved generally. From many types of media, a transmitter-receiver distance will be acquired, d_{ij}. Since this can be assumed to be real and positive, it can be squared as follows,

$$d_{ij}^2 = (\mathbf{r}_i - \mathbf{s}_j)^T(\mathbf{r}_i - \mathbf{s}_j) = \mathbf{r}_i^T\mathbf{r}_i + \mathbf{s}_j^T\mathbf{s}_j - 2\mathbf{r}_i^T\mathbf{s}_j. \tag{2.11}$$

The problem is then reformed according to the following invertible linear combinations of d_{ij}^2:

$$\mathbf{B} = \begin{pmatrix} d_{11}^2 & d_{12}^2 - d_{11}^2 & \dots & d_{1n}^2 - d_{11}^2 \\ d_{21}^2 - d_{11}^2 & & & \\ \dots & & \hat{\mathbf{B}} & \\ d_{m2}^2 - d_{11}^2 & & & \end{pmatrix}, \tag{2.12}$$

where the *compaction matrix* $\hat{\mathbf{B}}$ is an $(m-1) \times (n-1)$ matrix with entries as $\hat{B}_{ij} = \dfrac{d_{i,j}^2 - d_{i1}^2 - d_{1j}^2 + d_{11}^2}{-2}$, with $i = 2,\dots m$ and $j = 2,\dots,n$. The other elements in $\mathbf{B}$ are used as constraints for the solution.

The factorization can then be interpreted as follows. Let $\mathbf{R}_i = \big[(\mathbf{r}_i - \mathbf{r}_1)\big]$ and $\mathbf{S}_j = \big[(\mathbf{s}_j - \mathbf{s}_1)\big]$. Here $\hat{\mathbf{B}} = \mathbf{R}^T\mathbf{S}$ with $\mathbf{R}_i$ as columns of $\mathbf{R}$ and $\mathbf{S}_j$ as columns of $\mathbf{S}$. Since we assume that $\mathbf{R}$ and $\mathbf{S}$ are in a 3D affine space, the matrix $\hat{\mathbf{B}}$ has rank 3 at most. This also implies that in order to solve the problem, it is required that $m \geq 4$ and $n \geq 4$. By factorizing $\hat{\mathbf{B}}$, we can compute the vectors to all receivers and transmitters from unknown initial/reference positions ($\mathbf{r}_1$ and $\mathbf{s}_1$).

By fixing $\mathbf{r}_1$ at the origin and $\mathbf{s}_1$ as a vector from the origin, in terms of an affine transformation matrix $\mathbf{L}$ and vector $\mathbf{b}$, the problem is reformulated as follows,

$$\begin{aligned} &\mathbf{r}_1 = 0,\ \mathbf{s}_1 = \mathbf{L}\mathbf{b}, \mathbf{r}_i = \mathbf{L}^{-T}\tilde{\mathbf{R}}_i,\ i = 2 \dots m, \\ &\mathbf{s}_j = \mathbf{L}(\tilde{\mathbf{S}}_j + \mathbf{b}),\ j = 2 \dots n, \end{aligned} \tag{2.13}$$

where $\tilde{\mathbf{R}} = \mathbf{L}^T\mathbf{R}$, $\tilde{\mathbf{S}} = \mathbf{L}^{-1}\mathbf{S}$, and hence $\hat{\mathbf{B}} = \tilde{\mathbf{R}}^T\mathbf{L}^{-1}\mathbf{L}\tilde{\mathbf{S}} = \mathbf{R}^T\mathbf{S}$.

Using this parametrization, the equations from matrix $\mathbf{B}$, (2.46) become

$$\begin{aligned} d_{11}^2 &= (\mathbf{r}_1 - \mathbf{s}_1)^T(\mathbf{r}_1 - \mathbf{s}_1) = \mathbf{s}_1^T\mathbf{s}_1 = \mathbf{b}^T\mathbf{L}^T\mathbf{L}\mathbf{b} \\ &= \mathbf{b}^T\mathbf{H}^{-1}\mathbf{b}, \end{aligned} \tag{2.14}$$

$$\begin{aligned} d_{1j}^2 - d_{11}^2 &= \mathbf{s}_j^T\mathbf{s}_j - \mathbf{s}_1^T\mathbf{s}_1 = \tilde{\mathbf{S}}_j^T\mathbf{L}^T\mathbf{L}\tilde{\mathbf{S}}_j + 2\mathbf{b}^T\mathbf{L}^T\mathbf{L}\tilde{\mathbf{S}}_j \\ &= \tilde{\mathbf{S}}_j^T\mathbf{H}^{-1}\tilde{\mathbf{S}}_j + 2\mathbf{b}^T\mathbf{H}^{-1}\tilde{\mathbf{S}}_j, \end{aligned} \tag{2.15}$$

$$\begin{aligned} d_{i1}^2 - d_{11}^2 &= \mathbf{r}_i^T\mathbf{r}_i - 2\mathbf{r}_i^T\mathbf{s}_1 = \tilde{\mathbf{R}}_i^T(\mathbf{L}^T\mathbf{L})^{-1}\tilde{\mathbf{R}}_i - 2\mathbf{b}^T\tilde{\mathbf{R}}_i \\ &= \tilde{\mathbf{R}}_i^T\mathbf{H}\tilde{\mathbf{R}}_i - 2\mathbf{b}^T\tilde{\mathbf{R}}_i, \end{aligned} \tag{2.16}$$

where the symmetric matrix $\mathbf{H} = (\mathbf{L}^T\mathbf{L})^{-1}$. With this parameterization, there are in total 9 unknowns (6 and 3 unknowns for $\mathbf{H}$ and $\mathbf{b}$, respectively), and hence a solution can be found. Since this solution has its own coordinate system, with prior knowledge this can be transformed back to the original coordinate system.

3 Non-linear Optimization Approaches

In the development of the different systems for robust estimation, we use several different local optimization techniques. In particular we use methods for local optimization of the type

$$\min_{\mathbf{r},\mathbf{s}} \sum_{(i,j)\in\tilde{W}} f(d_{i,j} - ||\mathbf{r}_i - \mathbf{s}_j||_2), \tag{2.17}$$

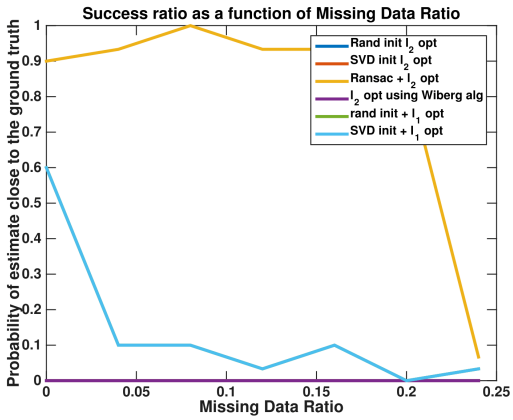
where $f(r)$ is chosen to be (i) $f(r) = r^2$ (l^2-norm), (ii) $f(r) = |r|$ (l^1-norm) or (iii) $f(r) = \min(r^2, T)$ (truncated l^2-norm). If the subset $\tilde{W}$ of the measurements contains no outliers and if the starting point is good, then the l^2-norm can give good estimates. Optimizing using the l^1-norm is less sensitive to the subset $\tilde{W}$ containing outliers, but still requires a reasonably good starting point to converge to a good solution. Local optimization of the truncated l^2-norm is even more sensitive to having a good starting point. Nevertheless, these local optimization techniques are important components for designing robust systems.

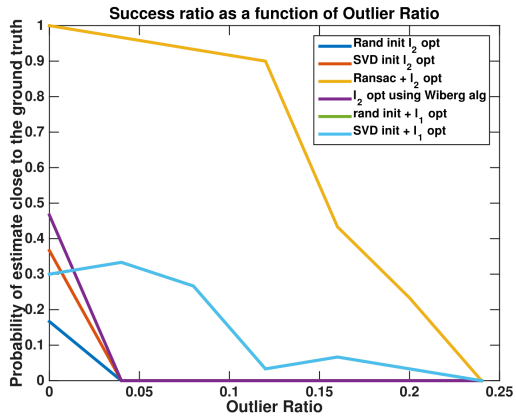
4 Obtaining Initial Estimates

Finding the optimal solution to problem 5, in the presence of outliers and missing data is a highly non-convex problem. We are thus dependent on finding good initial starting solutions, for the optimization methods from the previous section to work. We will in this section describe the different initialization methods that we have used in our experiment. In the next section we will describe our main contribution to the initialization problem.

Arguably, the most straight-forward way to initialize a solution, is to simply randomly place all receivers and senders within some space. This usually gives poor initial estimates, and the local optimization will be prone to get stuck in local minima. A slight improvement to this idea, is to use multiple restarts and optimize from each initial position, and then in the end choose the best solution.

Another way of initializing, that we have explored, is using the rank constraint on the compaction matrix. Here one can use many existing methods for doing the low rank matrix factorization. One important draw-back of these methods, is that we need to have at least one row and one column of the data matrix completely known, and without outliers. The last criteria is of course hard to check. If all data is known, the optimal low rank factorization is given by singular value decomposition (SVD) of the data matrix. A heuristic for handling missing data, is simply to fill in the missing data with some random values that follow the statistics of the other known measurements. One can then use SVD to obtain an initial estimate. This can be used directly to find the solution to the original problem as described in section 2. Alternatively, the initial low rank matrix factorization can be refined using the Wiberg algorithm, [16].





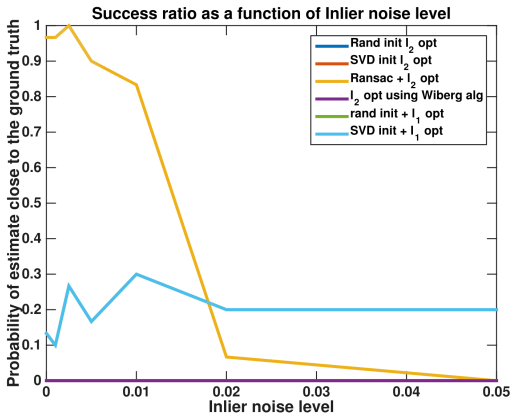


Figure 2.4: Comparison of success ratios for a number of tested systems, as functions of (from top left to bottom right) missing data ratio, outlier ratio and inlier noise level respectively.

5 Random Sampling Paradigm

The RANSAC or hypothesize and test paradigm, has proven to be useful in situations where there are outliers in the data, [14]. In this paradigm, a subset of the data is used to estimate the unknown parameters. The remainder of the data is then used to verify or falsify the parameters. This is typically repeated a fixed number of iterations. The parameters that give the largest number of inliers are then usually used as an initial estimate for the subsequent non-linear optimization of the parameters.

For Problem 5, there are several ways one could implement the hypothesis and test paradigm. One idea would be to use efficient algorithms for determining receiver and sender positions from minimal data, [6]. Although this solver and the test is relatively fast, we propose an alternative to this approach. The main idea is to find a fast way to hypothesize and test. We will use the rank constraints of the compaction matrix to do this. Our method is described in Algorithm 2.

Algorithm 2 Our RANSAC initialization scheme

1: Select 5 receivers randomly
2: Find all senders, for which there are no missing data to the 5 receivers
3: If there are at least five such senders, select 4 of these senders randomly
4: **Hypothesize:** Use the 5×4 matrix $\mathbf{F}$, with elements $\mathbf{F}_{i,j} = d_{i,j}^2$. Calculate the compaction matrix $\hat{\mathbf{B}}$ for $\mathbf{F}$ as shown in equation (2.46), which is a 4×3 matrix. Calculate a unit vector $\mathbf{v}$ which lies in the left null space of $\hat{\mathbf{B}}$.
5: **Test:** Assuming that a column $\mathbf{f}$, whose elements are $\mathbf{f}_i = d_{i,j_{test}}^2$ contain no outliers, then the vector $\hat{\mathbf{b}}$ is the compaction matrix of $[\mathbf{F}_1, \mathbf{f}]$ where $\mathbf{F}_1$ is the first column of $\mathbf{F}$. This then should have $\mathbf{v} \cdot \hat{\mathbf{b}} = 0$. Assuming low noise it is reasonable to declare it an inlier if $|\mathbf{v} \cdot \hat{\mathbf{b}}| < T$, where T is a threshold that depends on the noise level σ and the data d. Repeat this test for all the other columns.
6: Repeat steps 1-5, K times and keep track of the hypothesis that gave the largest number of inlier columns.

6 Experimental Evaluation

For the experimental evaluation, we generated a series of scenarios with different missing data ratio, outlier data ratio and different levels of inlier noise. For simplicity we have kept the number of receivers (m) and senders (n) fixed at $m = 30$ and $n = 30$. We also fixed the room size to be $10 \times 10 \times 3$ meters and placed the ground truth positions of the receivers and senders randomly in this box using a uniform distribution. Finally the errors $\epsilon_{i,j}$ for outliers were randomly drawn uniformly in the intervals $[-1.2, -0.4]$ and $[0.4, 1.2]$ meters. This means that for low levels of noise there is a relatively clear difference between

inlier and outlier distributions. For higher levels of noise the inlier and outlier distributions will start to overlap considerably. Notice, however, that it is difficult to determine which measurements are inliers by simply studying the distance measurements.

For each setting we generate a number of synthetic scenarios where the ground truth position of the receivers and senders are placed randomly as described above. A random subset W out of the mn measurements in I are chosen so that $|W|/|I| \approx 1 - \sigma_o$. Then a random subset $W_i \subset W$ of inlier measurements are chosen so that $|W_i|/|W| \approx 1 - \sigma_i$. The measurements $d_{i,j}$ are generated according to (2.91) with added noise according to the inlier model for $(i,j) \in W_i$ and according to the outlier model for $(i,j) \in W_o = W \setminus W_i$.

Each synthetic dataset d is then tested against each method. Here we tested the following systems

- **Ransac + l^2 opt:** Our proposed framework with initial estimate using RANSAC as described in Section 5, followed by l^2 optimization on detected inlier set.
- **SVD Init + l^1 opt:** Initial estimate using SVD as described in Section 4, followed by l^1 optimization.
- **Rand Init + l^1 opt:** Initial estimate using random placement as described in Section 4, followed by l^1 optimization
- **l^2 opt using Wiberg alg:** Initial estimate using SVD as described in Section 4, followed by rank 3 factorization using Wiberg algorithm followed by l^2 optimization.
- **SVD Init + l^2 opt:** Initial estimate using SVD as described in Section 4, followed by l^2 optimization

As can be seen in Figure 2.4, the proposed method outperforms the other methods in terms of robustness against outliers and missing data. For increased levels of noise the l^1 optimization methods degrades more gracefully.

For further experimental evaluation, our method was tested on real data measurements. In Figure 2.5, we conducted a round-trip time Wi-Fi experiment in 2D using 4 Nexus 6 phones as anchors and one phone as a transmitter. These phones come with a IEEE 802.11.mc Wi-Fi standard, which produces round-trip time measurements in metres. The experimental environment that was chosen was a large open space in an office block (Ideon Alfahuset, Lund, Sweden) with dimensions $\sim 12 \times 18$ m.

For the experiment, a ground truth was measured for all the anchor points and the path of the moving source every $0.5 \pm 0.005m$ with a tape measure. The experiment was conducted by walking the predefined path with the transmitter. In our experiment, our proposed

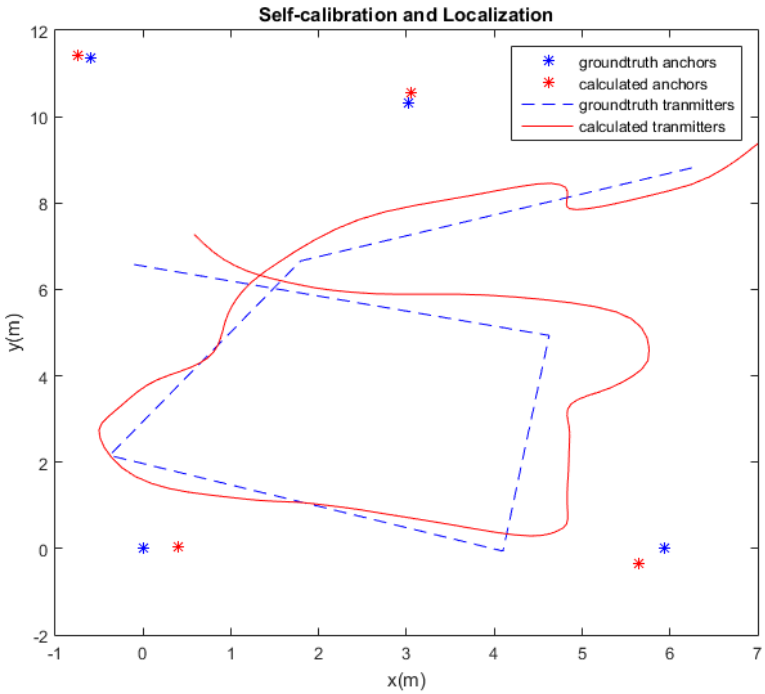


Figure 2.5: This figure illustrates the estimated anchor positions and the source positions. This is overlaid on the ground truth anchor positions and predefined path.

framework does not require prior knowledge about the positions for any of the anchors or source locations nor the number of anchors and source locations. The only requirement is that the number of anchors and source locations satisfy the requirement of the minimal solver, for this 2D case (3,3).

A fixed number of iterations was used; 20 iterations for the initial selection of 3 receivers and senders, then a further 140 iterations to extend the number of columns and rows. The tolerance was set to $T = 3$ for the initial selection and extension of columns then reduced to $T = 2$ for the extensions of the rows.

Once the initial values have been estimated, it undergoes l^2 optimization on the inlier set. We also added a smoothness prior in the optimization. This prior is based on minimizing acceleration, according to

$$res_a = \frac{1}{\sigma_a^2} \sum_{j=2}^{n-1} \|\mathbf{s}_{j-1} - 2\mathbf{s}_j + \mathbf{s}_{j+1}\|_2^2, \tag{2.18}$$

where σ_a is a parameter controlling the strength of the smoothness prior.

The ground truth anchor positions were $((0,0),(5.932,0),(3.0163,10.3079),(-0.6018,11.3658))$. The estimated anchors were calculated to be $((0.4044,0.0429),(5.6435,-0.3461),(3.0475,10.5527),(-0.7488,11.4242))$. This gives a total Euclidean distance error of 0.5923 m.

One section of the predefined path was obscured by a staircase. There the measurements had 6.52% missing data and 9.7826% of the data was considered to be outliers, according to our algorithm. In comparison to the whole data set we had 1.23% missing data and 3.3451% of the data was considered to be outliers. To further test our method, we conducted an experiment using ultra-wideband measurements in 3D with 6 anchors. Here we used the same number of iterations as before but using a tolerance of $T = 0.2$ and the 5 receivers and 5 senders 3D minimal solver, the result is shown in Figure 2.6.

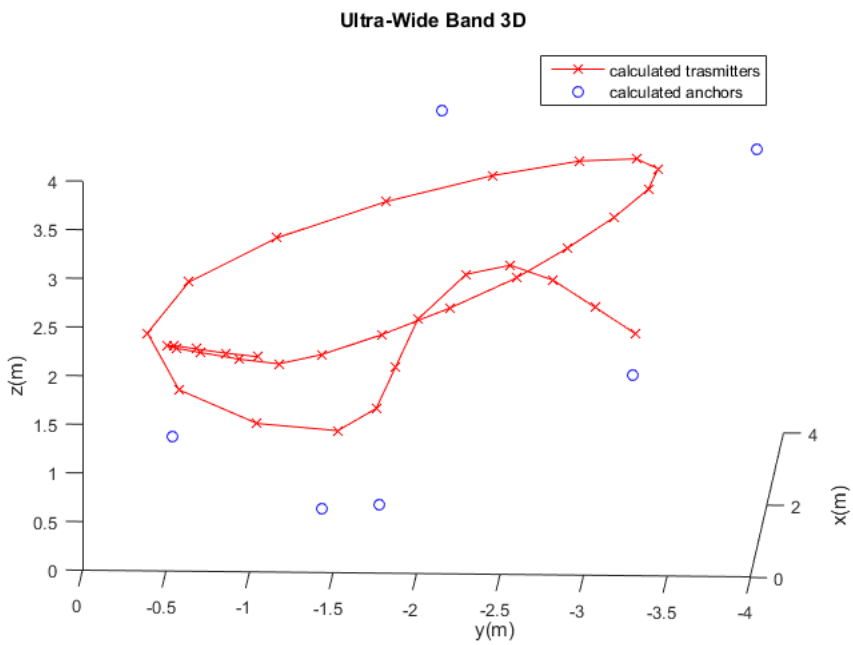


Figure 2.6: This figure illustrates the estimated anchor and source positions.

7 Conclusions

In this paper we have constructed several systems for estimating receiver-sender node positions from measured receiver-sender distances in the presence of outliers and missing data. We propose several new methods for solving these problem. In particular we propose a novel hypothesis and test framework that efficiently finds initial estimates of the unknown

parameters and combine such methods with efficient optimization techniques to obtain efficient, precise and robust systems. The proposed systems are evaluated against current state-of-the-art methods on a large set of benchmark tests. Our proposed hypothesis and test framework is then further tested on real Wi-Fi and ultra-wideband measurements to solve time-of-arrival self-calibration and localization.

Looking at Figure 2.5, we can see that a reasonable estimation was calculated. The general predefined path shape was preserved and the anchor positions are reasonable with a total error of 0.5923 m. Despite the path being obscured by stairs, the proposed method robustly and accurately estimated source positions. The experimental results demonstrates the robustness of the proposed method and how it can be effective in other areas as the errors in the estimations depend predominately on the accuracy of the distance measurements.

Bibliography

[1] N. Priyantha, H. Balakrishnan, E. Demaine, and S. Teller, "Anchor-free distributed localization in sensor networks," in *Proceedings of the 1st international conference on Embedded networked sensor systems*, pp. 340–341, ACM, 2003.

[2] G. Young and A. Householder, "Discussion of a set of points in terms of their mutual distances," *Psychometrika*, vol. 3, no. 1, pp. 19–22, 1938.

[3] P. Biswas, T.-C. Lian, T.-C. Wang, and Y. Ye, "Semidefinite programming based algorithms for sensor network localization," *ACM Trans. Sen. Netw.*, vol. 2, pp. 188–220, May 2006.

[4] T. Eren, O. Goldenberg, W. Whiteley, Y. Yang, A. Morse, B. Anderson, and P. Belhumeur, "Rigidity, computation, and randomization in network localization," in *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, vol. 4, pp. 2673–2684, IEEE, 2004.

[5] H. Stewénius, *Gröbner Basis Methods for Minimal Problems in Computer Vision*. PhD thesis, Lund University, APR 2005.

[6] Y. Kuang, S. Burgess, A. Torstensson, and K. Åström, "A complete characterization and solution to the microphone position self-calibration problem," in *The 38th International Conference on Acoustics, Speech, and Signal Processing*, 2013.

[7] S. Burgess, Y. Kuang, and K. Åström, "Node localization in unsynchronized time of arrival sensor networks," in *Proceedings of the 21st International Conference on Pattern Recognition*, 2012.

[8] S. Burgess, Y. Kuang, and K. Åström, "Pose estimation from minimal dual-receiver configurations," in *Proceedings of the 21st International Conference on Pattern Recognition*, 2012.

[9] M. Pollefeys and D. Nister, "Direct computation of sound and microphone locations from time-difference-of-arrival data," in *Proc. of International Conference on Acoustics, Speech and Signal Processing*, 2008.

[10] M. Crocco, A. Del Bue, M. Bustreo, and V. Murino, "A closed form solution to the microphone position self-calibration problem," in *37th International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2012), Kyoto, Japan*, March 2012.

[11] M. Crocco, A. Del Bue, and V. Murino, "A bilinear approach to the position self-calibration of multiple sensors," *Trans. Sig. Proc.*, vol. 60, pp. 660–673, feb 2012.

[12] S. Thrun, "Affine structure from sound," in *Proceedings of Conference on Neural Information Processing Systems (NIPS)*, (Cambridge, MA), MIT Press, 2005.

[13] Y. Kuang, E. Ask, S. Burgess, and K. Åström, "Understanding toa and tdoa network calibration using far field approximation as initial estimate," in *ICPRAM*, 2012.

[14] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–95, 1981.

[15] C. Eckart and G. Young, "The approximation of one matrix by another of lower rank," *Psychometrika*, vol. 1, no. 3, pp. 211–218, 1936.

[16] T. Wiberg, "Computation of principal components when data are missing," in *Proc. Second Symp. Computational Statistics*, 1976.

[17] H. Aanaes, R. Fisker, K. Åström, and J. Carstensen, "Robust factorization," *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2002.

[18] A. M. Buchanan and A. W. Fitzgibbon, "Damped newton algorithms for matrix factorization with missing data," in *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 2 - Volume 02*, CVPR '05, (Washington, DC, USA), pp. 316–322, IEEE Computer Society, 2005.

[19] Q. Ke and T. Kanade, "Robust L_1 norm factorization in the presence of outliers and missing data by alternative convex programming," in *Proc. Conf. Computer Vision and Pattern Recognition*, 2005.

[20] A. Eriksson and A. Hengel, "Efficient computation of robust weighted low-rank matrix approximations using the L_1 norm," *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2012.

[21] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?," *Journal of ACM*, 2009.

[22] R. Garg, A. Roussos, and L. Agapito, "Dense variational reconstruction of non-rigid surfaces from monocular video," in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pp. 1272–1279, IEEE, 2013.

[23] C. Olsson and M. Oskarsson, "A convex approach to low rank matrix approximation with missing data," in *Scandinavian Conf. on Image Analysis*, 2011.

[24] R. Cabral, F. D. la Torre, J. P. Costeira, and A. Bernardino, "Unifying nuclear norm and bilinear factorization approaches for low-rank matrix decomposition," in *Proc. Int. Conf. on Computer Vision*, 2013.

[25] L. W. Mackey, M. I. Jordan, and A. Talwalkar, "Divide-and-conquer matrix factorization," in *Advances in Neural Information Processing Systems*, pp. 1134–1142, 2011.

[26] V. Larsson, C. Olsson, E. Bylow, and F. Kahl, "Rank minimization with structured data patterns," in *ECCV*, 2014.

[27] F. Jiang, M. Oskarsson, and K. Åström, "On the minimal problems of low-rank matrix factorization," in *Proc. Conf. Computer Vision and Pattern Recognition*, 2015.

Paper III

Robust Self-Calibration of Constant Offset Time-Difference-of-Arrival

Kenneth Batstone[1],Gabrielle Flood[1], Thejasvi Beleyur[2],Viktor Larsson[3], Magnus Oskarsson[1], Holger R. Goerlitz Kalle[3] Åström[1]

[1] Centre of Mathematical Sciences, Lund University, Lund Sweden

[2] Dept. of Computer Science, ETH Zurich, Switzerland

[3] Acoustic and Functional Ecology, Max Planck Inst. for Ornithology, Seewiesen, Germany

Abstract: In this paper we study the problem of estimating receiver and sender positions from time-difference-of-arrival measurements, assuming an unknown constant time-difference-of-arrival offset. This problem is relevant for example for repetitive sound events. In this paper it is shown that there are three minimal cases to the problem. One of these (the five receiver, five sender problem) is of particular importance. A fast solver (with run-time under 4 μs) is given. We show how this solver can be used in robust estimation algorithms, based on RANSAC, for obtaining an initial estimate followed by local optimization using a robust error norm. The system is verified on both real and synthetic data.

1 Introduction

The problem of estimating receiver-sender node positions from measured arrival times of radio or sound signals is a key issue in different applications such as microphone array calibration, radio antenna array calibration, mapping and positioning. This field is well researched but in this paper we will focus on the anchor-free sensor network calibration both in terms of time-of-arrival measurements (TOA) and time-difference-of-arrival measurements (TDOA). For time-of-arrival the planar case of three receivers and three senders (3R/3S) was solved in [1]. For the full 3D case the over-determined problem (10R/4S) was studied in [2], where a solver for this non-minimal case was provided. There are actually three minimal cases for the 3D case, namely (4R/6S), (5R/5S) and (6R/5S). A practical solver was presented in [3]. There are in general 38, 42 and 38 solutions respectively for the three different set ups. Faster solvers for these minimal cases were provided in [4].

In this paper we study the constant offset TDOA self-calibration problem. It is a problem that naturally arises e.g. when signals are emitted with a known period. As an estimation problem it lies between TOA and full TDOA. In the paper we study the minimal (5R/5S) problem and provide a fast (few μs) solver. Robust parameter estimation often use the hypothesize and test paradigm, e.g. using random sampling consensus, [5] or one of its many variants [6, 7, 8]. In these frameworks minimal solvers are important building blocks for generating model hypotheses, and we show in the paper how a minimal solver can be used for robust parameter estimation of sender positions, receiver positions and unknown offset. The system is capable of handling missing data, outliers and noise. The algorithms are tested on synthetic data as well as real data, in an office environment and in a cave. The methods are straightforward to generalize for degenerate configurations which arise if senders or receivers are restricted to a plane or to a line.

2 Time-Difference-of-Arrival Self Calibration

The problem we are considering involves m receiver positions $\mathbf{r}_i \in \mathbb{R}^3$, $i = 1, \ldots, m$, and n sender positions $\mathbf{s}_j \in \mathbb{R}^3$, $j = 1, \ldots, n$. This could for example represent the microphone positions and locations of sound emissions, respectively. Assume that the arrival time of a sound j to receiver i is t_{ij} and that the time that sound j is emitted is T_j. Multiplying the travel time $t_{ij} - T_j$ with the speed v of the signal we obtain the distance between senders and receiver,

$$v(t_{ij} - T_j) = \|\mathbf{r}_i - \mathbf{s}_j\|_2, \tag{2.19}$$

where $\|.\|_2$ is the l^2-norm. The speed v is throughout the paper assumed to be known and constant.

In many settings the times of emissions T_j are unknown, but regular, *e.g.*

$$T_j = k_1 j + k_0, \tag{2.20}$$

where the interval k_1 is known. Inserting (2.20) into (2.19) we obtain

$$v(t_{ij} - k_1 j - k_0) = \|\mathbf{r}_i - \mathbf{s}_j\|_2. \tag{2.21}$$

Assuming an erroneous (but regular) emission time $\tilde{T}_j = k_1 j + \tilde{k}_0$ and introducing (the measured) $z_{ij} = v(t_{ij} - \tilde{T}_j)$ and (the unknown) $o = v(k_0 - \tilde{k}_0)$ yields the following expression

$$z_{ij} = \|\mathbf{r}_i - \mathbf{s}_j\|_2 + o. \tag{2.22}$$

Note that this is a simplified variant of the general time-difference-of-arrival problem (see *e.g.* [9]), which allows for a different offset o for every j,

$$z_{ij} = \|\mathbf{r}_i - \mathbf{s}_j\|_2 + o_j. \tag{2.23}$$

Problem 3. *(Constant Offset Time-Difference-of-Arrival Self-Calibration) Given measurements* $\tilde{z}_{ij}$

$$\tilde{z}_{ij} = \|\mathbf{r}_i - \mathbf{s}_j\|_2 + o + \epsilon_{ij}, \tag{2.24}$$

for a subset $W \subset I$ *of all the receiver-sender index pairs* $I = \{(i,j)|i = 1,\dots m, j = 1,\dots,n\}$ *determine receiver positions* $\mathbf{r}_i$, $i = 1,\dots,m$ *and sender positions* $\mathbf{s}_j$, $j = 1,\dots,n$ *and offset* o. *Here the errors* ϵ_{ij} *are assumed to be either* **inliers**, *in which case the errors are small (*$\epsilon_{ij} \in N(0,\sigma)$*) or* **outliers**, *in which case the measurements are way off.*

Here we will use the set W_{in} for the indices (i,j) corresponding to the inlier measurements and W_{out} for the indices corresponding to the outlier set.

3 Local Optimization and the Low Rank Relaxation

If an initial estimate of the parameters $\theta_1 = \{R, S, o\}$ is given and if the set of inliers is known, then refinement of the estimate can be found by optimization methods, *e.g.* Levenberg-Marquardt (LM) [10, 11],

$$\min_{\theta_1} f(\theta_1) = \sum_{(i,j)\in W_{\text{in}}} (z_{ij} - (\|\mathbf{r}_i - \mathbf{s}_j\|_2 + o))^2. \tag{2.25}$$

There is an interesting relaxation to the problem, that exploits the fact that the matrix with elements $(z_{ij}-o)^2$ is rank 5, [2]. Further simplifications use the double compaction method [9]. The double compaction matrix M is defined as the matrix with elements

$$M_{ij} = (z_{ij} - o)^2 - a_i - b_j, \tag{2.26}$$

and it can be shown to have rank 3, i.e. $M = U^T V$, where U is of size $3 \times m$ and V is of size $3 \times n$. The relaxed problem involves a set of parameters $\theta_2 = \{U, V, b, a, o\}$. Here the constraints can be written as

$$z_{ij} = \sqrt{u_i^T v_j + a_i + b_j} + o, \tag{2.27}$$

where u_i denotes column i of U and v_j denotes column j of V. Refinement of parameters can be done by performing local optimization on

$$\min_{\theta_2} f(\theta_2) = \sum_{(i,j)\in W_{\text{in}}} \left(z_{ij} - (\sqrt{u_i^T v_j + a_i + b_j} + o) \right)^2. \tag{2.28}$$

4 Minimal Problems and Solvers

By counting equations and unknowns, one finds that there are three minimal problems. The first two are the symmetric case when $m = 4, n = 7$ or $m = 7, n = 4$. This case is not addressed in this paper, but we believe it to be difficult to solve. The other case is $m = n = 5$. Here, we first present a solver for the constant offset and then discuss how to solve for sender and receiver positions.

Given a 5×5 matrix, Z, with time-difference-of-arrival measurements z_{ij}, the rank 3 constraint on the double compaction matrix in (2.26) can be written as

$$f(o) = \det(C^T (Z - o)^{\circ 2} C) = 0, \tag{2.29}$$

where

$$C = \begin{pmatrix} -1 & -1 & -1 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.30}$$

and $^{\circ 2}$ denotes element-wise squaring (Hadamard power). Although the elements of $(Z - o)^{\circ 2}$ are of degree 2 in o, the quadratic terms cancel out after multiplication with C^T and C. Thus the elements of $C^T (Z - o)^{\circ 2} C$ are linear in o. Since the determinant is linear in each column, the determinant $f(o)$ is a polynomial of degree four in the offset o. This can be summarized as

Theorem 1. *Given time-difference-of-arrival measurements from five receivers to five senders, there are four possible offsets o, given as the roots to the fourth degree polynomial $f(o)$, counting complex roots and multiplicity of roots.*

Table 2.1: Execution times for 5×5 minimal solvers steps. Notice that the steps of calculating o and the relaxed solution is significantly faster than upgrading to the full solution

Implementation	Matlab	C++
Calculation of o	38 μs	3.7 μs
Calculation of $\theta_2 = \{U, V, a, b, o\}$	100 μs	N/A
Calculation of $\theta_1 = \{R, S, o\}$	600 ms	22 ms

For each solution o it is possible to generate a solution θ_2 to the relaxed problem, according to

$$b = \begin{pmatrix} (z_{11}-o)^2 & (z_{12}-o)^2 & (z_{13}-o)^2 & (z_{14}-o)^2 & (z_{15}-o)^2 \end{pmatrix},$$

$$a = \begin{pmatrix} 0 \\ (z_{21}-o)^2 - (z_{11}-o)^2 \\ (z_{31}-o)^2 - (z_{11}-o)^2 \\ (z_{41}-o)^2 - (z_{11}-o)^2 \\ (z_{51}-o)^2 - (z_{11}-o)^2 \end{pmatrix}, \tag{2.31}$$

$$U = \begin{pmatrix} 0 & u_2 & u_3 & u_4 & u_5 \end{pmatrix}, \tag{2.32}$$

$$V = \begin{pmatrix} 0 & v_2 & v_3 & v_4 & v_5 \end{pmatrix}, \tag{2.33}$$

where $\begin{pmatrix} u_2 & u_3 & u_4 & u_5 \end{pmatrix}^T \begin{pmatrix} v_2 & v_3 & v_4 & v_5 \end{pmatrix}$ is any rank 3 factorization of the matrix $C^T(Z - o)^{\circ 2}C$.

From a solution θ_2 to the relaxed problem it is possible to upgrade to a solution θ_1 to the original problem. This involves solving a system of polynomial equations. The procedure was first described in [3], where an algorithm for solving this was presented. Recently, a faster algorithm was presented in [4].

An efficient implementation for calculating the four solutions of the offset o given the measurements z takes 4 μs for a C++-implementation. Generating the solution θ_2 to the relaxed problem adds a few μs. However, calculating a solution θ_1 to the original problem takes another 22 ms. Thus, it is advantageous to estimate the parameters of the relaxed problem and postpone the upgrade from θ_2 to θ_1 as a final step, see Table 2.1.

5 Using RANSAC for Five Rows

We propose the use of the fast minimal solver in an hypothesize and test framework to obtain (i) a initial estimate on the offset o and (ii) an initial inlier set. The steps are described in Algorithm 3

Algorithm 3 Offset RANSAC

1: Randomly select 5 rows and columns. Find the four solutions on o given the time-difference-of-arrival measurements.
2: For each solution o, calculate the relaxed solution $\theta_2 = \{U, V, a, b, o\}$.
3: For selected rows and for each remaining column, check for inliers according to the residuals in (2.28).

6 Robust Estimation of Parameters

We use these minimal solvers with RANSAC as described in the previous section to find one or several initial estimates of the parameters θ_2 for a subset of five receivers and k senders. The solution is extended to additional rows and/or columns using robust techniques as described in [12]. During this process it is useful to keep the errors down by occasionally refining the solutions using local optimization. This has shown to reduce failures, see e.g. [13, 14]. In the proposed estimation algorithm we postpone the upgrade from θ_2 to θ_1 until we have found a good solution involving a large portion of the receiver and sender positions.

7 Experimental Validation

7.1 Minimal Solver

To test the numerical accuracy and robustness of our minimal solver we conducted an experiment using simulated data without noise. We generated a large number of instance problems (10,000) with known offsets. We then ran our solvers and compared the returned solutions with the ground truth solution. For each instance problem we recorded the distance to the closest solution. In Figure 2.7 the resulting histogram of the logarithm of the absolute errors are shown. As can be seen, both implementations get close to machine precision.

7.2 Experimental Setup for Real Data

We have tested our system on (i) experiments made in an office environment and (ii) experiments made at the Orlova Chuka cave, Bulgaria.

For the office experiments, 12 microphones (8x t.bone MM-1, 4x Shure SV100) were positioned around a room ($\sim 3 \times 5\ m^2$) and measured using a laser to obtain ground truth

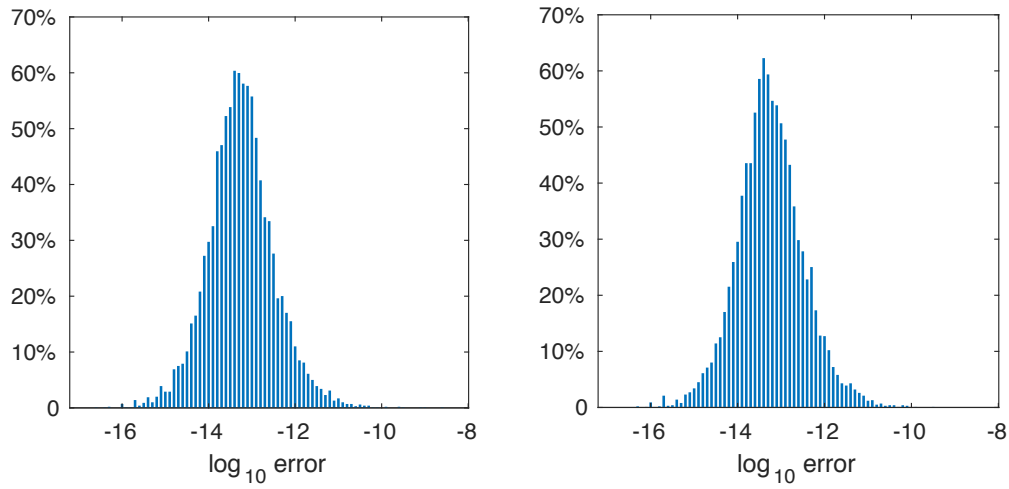


Figure 2.7: Left shows the histogram of the logarithm of the absolute errors, for the Matlab implementation of our minimal solver. To the right the corresponding histogram for the C++ implementation.

positions of the microphones with an error of ± 2 *mm*. The space was cleared of most the furniture to create an open space to conduct the experiment in. The sound recordings were captured using a Roland UA-1610 Sound Capture audio interface and automatically amplified. The recordings were made using the open source software Audacity 2.3.0 with a sampling frequency of 96 *kHz* on a laptop. A synthetically generated chirp was then played using a simple loudspeaker every half second for 30 *s* while moving the speaker around in the room.

For the cave experiments, 12 microphones (4x Sanken CO-100K, 8x Knowles SPU0410) were positioned in a section of the cave, four microphones were placed on an inverted T array near one wall, while the other eight microphones were placed on the adjacent wall. The sound recordings were captured using pre-amplifiers (Quadmic, RME) and two synchronised Fireface 800 (RME) audio interfaces running at a sampling frequency of 192 *kHz*. Recording and playback were controlled via a custom written script based on the sound device library [15] in Python 2.7.12 [16]. Ultrasonic chirps (8 *ms*, $16 - 96$ *kHz* upward hyperbolic sweep) were played every second via one of the audio interfaces, amplified (Basetech AP-2100) and presented through a Peerless XT25SC90-04 loudspeaker. The speaker was attached to a 3-m-long pole and slowly waved in the approximately $5 \times 9 \times 3$ m^3 recording volume. Playbacks were done past 6:00 am to prevent disturbing the resident bat population.

7.3 Experimental Evaluation for Real Data

Once the office recordings were taken, an algorithm was used to find the chirps in the captured sound recordings and the algorithm then outputs the z_{ij} matrix. This can then be used in our RANSAC scheme, Algorithm 3. For this experiment we used the (5R/5S) minimal solver. A fixed number of iterations was used; 100 iterations for the initial selection of 5 receivers and senders, then the extension to more columns and rows was allowed until there was no better solution. The tolerance was set to $T = 0.01$ for the initial selection and extension of rows and column.

Once the initial values have been estimated, it underwent l^2 optimization on the inlier set. The results of the estimated microphone positions after the optimization are shown in Figure 2.8.

This produced an Euclidean distance error between each of the microphones calculated position and its ground truth position as $(0.2016, 0.0587, 0.1444, 0.1153, 0.2017, 0.1326, 0.1407, 0.1198, 0.2041, 0.2010, 0.1908, 0.2110)$ m.

For graphical purposes, a Procrustes fitting was used on the microphone positions to spread the total error over all 12 microphones. In the Procrustes fitting only rotation and translation were allowed.

For the cave experiment a similar scheme was devised and the results are shown in Figure 2.9.

8 Conclusions

In this paper, a novel method has been constructed to efficiently solve a TDOA problem with a constant offset. This has been verified using simulated data to test the solver and real experimental data to test our algorithms in realistic scenarios.

Looking at Figure 2.7 and Table 2.1, it can be seen that the calculation of the offsets and the calculation of the relaxed form θ_2 are very fast solvers without loss in numerical accuracy. The advantage of this is that when using a RANSAC approach, the iterations are performed quickly, giving a good initial estimate in which to optimize over, which is important in highly non-linear systems such as this.

Looking at the results from the office experiment, Figure 2.8, we can see that the calculated microphone positions are accurate and the residuals are small, mostly in the range ± 0.04 m. Further to this our inlier set appears to be accurate. The first and last few columns (corresponding to sound emissions) are not used in our initialisation. This is correct because the recording started before the chirps were sounded and ended after, so the chirp detection

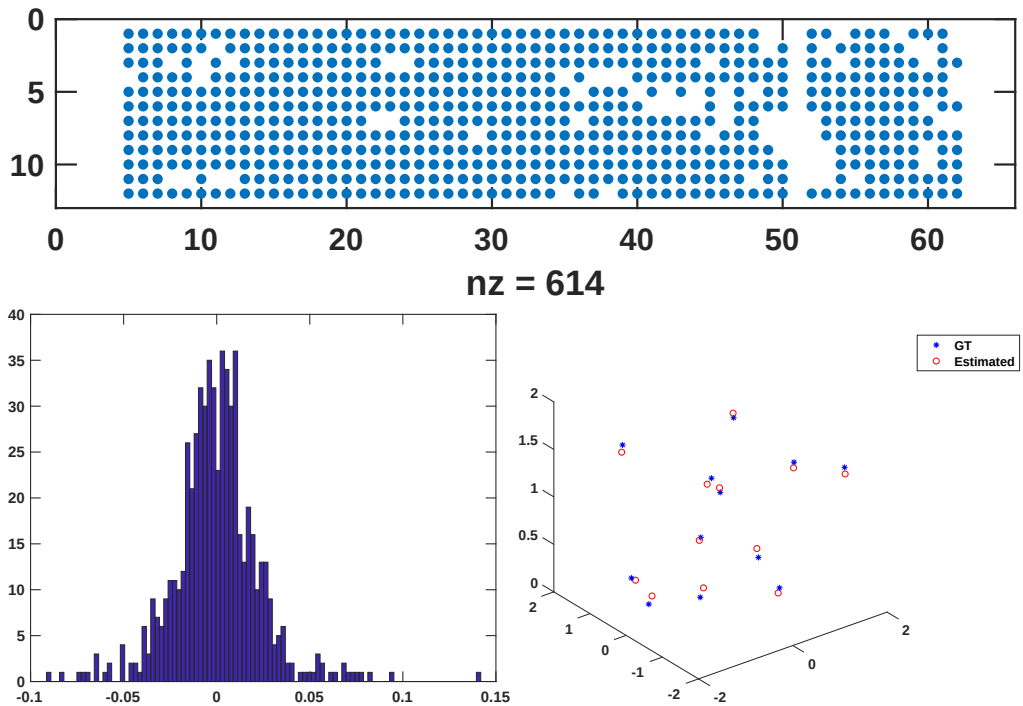


Figure 2.8: For the office experiment the figure shows detected inliers W_{in} (top), inlier residual histogram (bottom left), and estimated and ground truth microphone positions (bottom right).

algorithm falsely determined that they were also chirps but our method decided that the data in those regions do not fit the model. A comparison of the calculated microphone positions were made to a solution from a Full TDOA system, [9], which produced similar results and very similar residuals. This provided a sanity check that the chirp detection was working correctly and that from this dataset a better solution could not be found.

For the cave experiment, similar conclusions can be made, since the residuals are very low, we can conclude that we have an accurate model. This gives a real life example of how algorithms such as the one proposed can be used.

For future work, the study of the number of inliers could be of use. At the moment our algorithm may not extend to more rows and columns if the initial solution is poor, perturbing our final solution. Perhaps a method which could adapt the initial selection in order to give a required amount of inliers could be more advantageous.

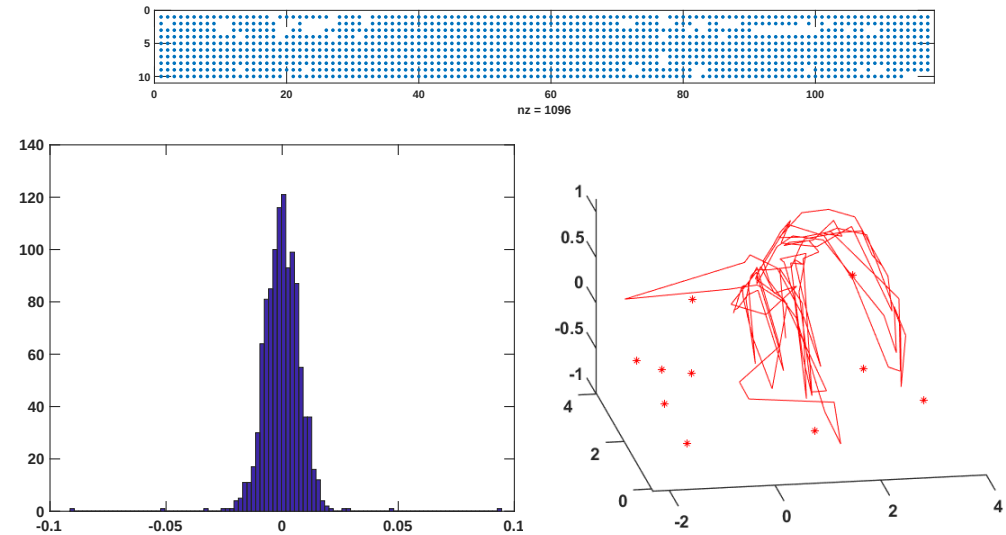


Figure 2.9: For the cave experiment the figure shows detected inliers $\mathcal{W}_{\text{in}}$ (top), inlier residual histogram (bottom left) and estimated microphone and sound source positions, red dots and line respectively (bottom right).

Bibliography

[1] H. Stewénius, *Gröbner Basis Methods for Minimal Problems in Computer Vision*. PhD thesis, Lund University, APR 2005.

[2] M. Pollefeys and D. Nister, "Direct computation of sound and microphone locations from time-difference-of-arrival data," in *Proc. of International Conference on Acoustics, Speech and Signal Processing*, 2008.

[3] Y. Kuang, S. Burgess, A. Torstensson, and K. Åström, "A complete characterization and solution to the microphone position self-calibration problem," in *The 38th International Conference on Acoustics, Speech, and Signal Processing*, 2013.

[4] V. Larsson, K. Åström, and M. Oskarsson, "Polynomial solvers for saturated ideals," in *International Conference on Computer Vision (ICCV)*, IEEE, 2017.

[5] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–95, 1981.

[6] O. Chum, J. Matas, and J. Kittler, "Locally optimized ransac," in *Joint Pattern Recognition Symposium*, pp. 236–243, Springer, 2003.

[7] R. Raguram, O. Chum, M. Pollefeys, J. Matas, and J.-M. Frahm, "Usac: a universal framework for random sample consensus.," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 2022–2038, 2013.

[8] S. Korman and R. Litman, "Latent ransac," in *Proc. Conf. Computer Vision and Pattern Recognition*, pp. 6693–6702, 2018.

[9] Y. Kuang and K. Åström, "Stratified sensor network self-calibration from tdoa measurements," in *Signal Processing Conference (EUSIPCO), 2013 Proceedings of the 21st European*, pp. 1–5, IEEE, 2013.

[10] K. Levenberg, "A method for the solution of certain non-linear problems in least squares," *Quarterly of applied mathematics*, vol. 2, no. 2, pp. 164–168, 1944.

[11] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *Journal of the society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963.

[12] K. Batstone, M. Oskarsson, and K. Åström, "Robust time-of-arrival self calibration with missing data and outliers," in *Signal Processing Conference (EUSIPCO), 2016 24th European*, pp. 2370–2374, IEEE, 2016.

[13] C. Engels, H. Stewenius, and D. Nister, "Bundle adjustment rules," in *PCV06*, 2006.

[14] G. Klein and D. Murray, "Parallel tracking and mapping for small ar workspaces," in *Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium on*, pp. 225–234, IEEE, 2007.

[15] M. Geier, "sounddevice 0.3.5." `https://python-sounddevice.readthedocs.io/en/0.3.5/_modules/sounddevice.html`, 2015.

[16] G. Van Rossum and F. L. Drake Jr, *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam, 1995.

Paper IV

Trust No One: Low Rank Matrix Factorization Using Hierarchical RANSAC

Magnus Oskarsson, Kenneth Batstone and Kalle Åström
Centre of Mathematical Sciences, Lund University, Lund, Sweden

Abstract: In this paper we present a system for performing low rank matrix factorization. Low-rank matrix factorization is an essential problem in many areas, including computer vision with applications in affine structure-from-motion, photometric stereo, and non-rigid structure from motion. We specifically target structured data patterns, with outliers and large amounts of missing data. Using recently developed characterizations of minimal solutions to matrix factorization problems with missing data, we show how these can be used as building blocks in a hierarchical system that performs bootstrapping on all levels. This gives a robust and fast system, with state-of-the-art performance.

1 Introduction

We will in this paper address the problem of robust estimation of low rank factorizations of matrices with missing data and outliers. Many problems in geometric computer vision can be formulated as such. Two examples are

- Affine structure-from-motion (SfM), where the observation matrix containing feature tracks can be factorized into the camera motions and the 3D structure.
- Photometric stereo, where the directions of light sources and the surface normals are separated by factorizing the measurement matrix composed of pixel intensities under a Lambertian model.

Other applications can be found in [1, 2, 3, 4]. These problems often lead to measurement matrices with highly structured data, in terms of which measurements that are available. In this paper we specifically target problems that exhibit such structured data patterns. Without missing data, the optimal solution to low rank matrix factorization, under the l^2-norm, is given by truncating the singular value decomposition of the measurement matrix, see [5]. When there are measurements missing in the data, there is no closed form solution to the l^2-norm minimization problem. The Wiberg algorithm [6] was the first method to handle missing data. A modified version of the Wiberg algorithm was presented in [7]. In [8], a damped Newton method is proposed to handle the missing data. If there are gross outliers in the data, optimizing the l^2-norm can give poor results. In [9], Aanaes *et al.* proposed an iteratively re-weighted least squares approach to optimize the objective function for robustness to outliers. Using more robust norms to outliers was considered in [10], where algorithms based on alternating optimization under the Huber-norm and the l^1-norm were introduced. Eriksson and Hengel generalized the Wiberg algorithm in [11], to using the l^1-norm. Sometimes, extra constraints can be posed on the factorization matrices. In [12, 13] constraints that the solution should lie on a certain manifold are considered and incorporated in the formulation. Due to the non-convexity of the matrix factorization, most methods mentioned above are based on alternating optimization, and are prone to get trapped in local minima. To address this issue, several works, such as [14, 15, 16, 3, 17] reformulate the problem to minimize the convex surrogate of the rank function – the nuclear norm. This makes it possible to use convex optimization to find the global optimum of the approximated objective function. These approaches can handle the problems when the rank is not known a priori. However, for applications with a given rank, the nuclear norm based methods usually perform inferior to the bilinear formulation-based methods [18]. The convex formulations often have problems with very high amounts of missing data and outliers. A way of handling highly structured data matrices is to divide the whole matrix into overlapping sub-blocks and combine the sub-block solutions, see [19, 20, 21, 22, 23]. Most of these methods do not consider both outliers and missing data at the same time.

There are a number of works that target specific computer vision applications for incomplete data. Examples are relative orientation problems, [24, 25], batch structure from motion estimation, [26], multi-dimensional scaling, [27], and shape estimation, [28]. It has also been shown that the specific problem of structure from motion with missing data is NP-hard, [29].

In this paper we largely build upon the work in [23] where minimal solvers for low rank factorizations of matrices with missing data were introduced. The emphasis was on how to analyze, describe and solve minimal problems. In this paper we address a number of algorithmic challenges (speed, accuracy, avoidance of local minima, robustness to outliers). Our contribution is a system that estimates a low rank factorization of a measurement matrix with, large amounts of missing data, in highly structured data patterns. It is based on bootstrapping minimal solvers, which gives speed and robustness to outliers. Running the solvers in a hierarchical manner gives tractable behaviour for larger input matrices, and easy parallelization. The system makes it possible to add additional constraints on the solution, throughout the pipeline.

2 Problem Formulation

We will look at the problem of finding a low rank matrix approximation to a given matrix X. Any rank K matrix X of size $M \times N$ can be written as $UV = X$, where U is a rank K matrix of size $M \times K$ and V is a rank K matrix of size $K \times N$. If X represents measurements of something, that in theory should have rank K, we can not expect this equation to hold exactly due to noise. We could then instead look at the norm of the residual between the measurement matrix X and the model UV, *i.e.*

$$e = \|X - UV\|_F. \tag{2.34}$$

In many cases one does not have access to all measurements, *i.e.* not all entries of X are known. We can then represent which measurements are known by the *index matrix* W of size $M \times N$, where the entries are either zero, if the corresponding measurement is unknown, or one if the corresponding measurement is known. We can then write the corresponding residual norm as,

$$e = \|(X - UV) \odot W\|_F, \tag{2.35}$$

where $\odot$ represents element-wise multiplication. In addition to measurement noise, one can also have gross outliers in the data, in some applications. In this case minimizing an error norm based on an l^2-distance often gives bad results. In order to decrease the influence of the outliers, robust norms are used, such as the l^1-norm or the Huber norm. In [30] a more refined loss function is proposed. If we assume that the inlier residuals approximately follow a Gaussian distribution, whereas outlier residuals have approximately

uniformly distributed errors, then this leads to the loss function

$$l(r) = -\log(c + \exp(-r^2)), \tag{2.36}$$

where r is the residual error. Truncating the squared error

$$l(r) = \begin{cases} r^2 & \text{if } |r| \leq \epsilon \\ \epsilon^2 & \text{otherwise} \end{cases} \tag{2.37}$$

gives a good approximation. If we denote the residual matrix $R = (X - UV) \odot W$, with entries r_{ij}, we can formulate our problem as

$$\underset{U,V}{\text{minimize}} \quad \sum_{i,j} l(r_{ij}). \tag{2.38}$$

The final error depends on ϵ which we set as a parameter in our algorithm. This is the bound that differentiates an inlier from an outlier measurement.

3 Matrix Factorization with Missing Data

We will use the characterization of low rank matrix factorization problems that was described in [23]. For completeness and readability, we will in short describe some of the results from that paper.

A key question is to study for which index matrices W of type $M \times N$, the problem of calculating the rank K matrix $X = UV$ is minimal and well-defined. For this we introduce the manifold $\Omega = R^{M\times K} \times R^{K\times N}$ of possible solutions $z = (U, V)$. The solution set $\Omega_{X,W} \subset \Omega$ for a given data matrix X and a index matrix W is defined as

$$\Omega_{X,W} = \{z = (U, V) \in \Omega | W \odot (X - UV) = 0\}. \tag{2.39}$$

Typically if the index matrix W has to many non-zero elements, then there are too many constraints in $W \odot (X - UV) = 0$ and the solution set $\Omega_{X,W}$ is empty for general X. If there is a solution $X = UV$, then we are interested to known if the solution is unique up to the so called *Gauge freedom* $z = (U, V)$ vs $z = (UH, H^{-1}V)$, where H is a general invertible $K \times K$ matrix.

Assume that input data matrix $X = U_0 V_0$ of size $M \times N$ has been generated by multiplying matrices U_0 of size $K \times M$ and V_0 of size $K \times N$. Assume also that both of these matrices are general in the sense that all $K \times K$ submatrices of both U_0 and V_0 have rank K. Furthermore assume that $M \geq K$ and $N \geq K$. An index matrix W is said to be *rigid*,

if the solution set $\Omega_{X,W}$ locally around the point $z_0 = (U_0, V_0)$ only consists of the set $z(H) = \{(U_0 H, H^{-1} V_0) | H \text{ invertible } K \times K \text{ matrix}\}$.

Since we are assuming that every sub-minor of U_0 and V_0 has full rank, one may actually fix the Gauge freedom by keeping one such sub-minor fixed. For example we could study the solutions for the points $z = (U, V)$ such that the first K rows of U are equal to those of U_0.

For two index matrices W_1 and W_2 of the same size we say that $W_1 \leq W_2$ if the inequality holds for every element. We say that $W_1 < W_2$ if $W_1 \leq W_2$ and $W_1 \neq W_2$. It is trivial to see that if W is rigid and if $W \leq W'$ then W' is also rigid. It also can be shown that if W' is rigid and overdetermined, then there is at least one $W < W'$ that is rigid and minimal. We say that an index matrix W is *minimal* if it is rigid and satisfies $\sum_{ij} W(i,j) = MK+NK-K^2$. For a minimal index matrix W and for general data X the solution set $\Omega_{X,W}$ consists of a finite number of points n_W up to the gauge freedom.

3.1 Henneberg Extensions

We will now describe how to generate the minimal problems. The inspiration comes from rigid graph theory, where the Henneberg construction is used to generate the Laman graph, see [31, 32]. The idea is that one starts with the smallest minimal index matrix, and by a series of extensions generate every minimal index matrix. For the rank K problem the smallest index matrix is a matrix of size $K \times K$ consisting of ones only.

There are both *constructive* extensions and *non-constructive* extensions. For a constructive extension from W to W', one can infer the number of solutions $n_{W'}$ from n_W and construct the solver, denoted by $f_{W'}$ from f_W. For non-constructive extensions, it can be shown that W is minimal if and only if W' is minimal. However, we can in general neither infer the number of solutions $n_{W'}$ from n_W nor derive a solver $f_{W'}$ from f_W. Certain of these constructive extensions are particularly fast and efficient. The simplest one is as follows.

Given a minimal index matrix W for a rank-K problem of size $M \times N$, an extended minimal index matrix W' of size $M \times (N+1)$ is formed by adding a column with exactly K elements set to one. The number of solutions are identical, i.e. $n_W = n_{W'}$. Extending an algorithm from f_W to $f_{W'}$ is straightforward. A similar extension can be done by adding a row with K indices.

3.2 Henneberg Reductions

There is also a simple recursive method to check if an index matrix W can be generated using only Henneberg 1 extensions. The procedure is as follows. Start with an index matrix

of size $M \times N$. If $M = N = K$ then the index matrix is minimal if and only if the matrix consists only of ones. If M or N is larger than K, we calculate the minimal number of ones for a row or column. If this number is less than K, then it can be shown that the index set in question is non-rigid. If this number is larger than K it can be shown that the index set (if minimal) cannot be generated by Henneberg 1 extensions only. Finally if the number is K, then we can remove the row (or column) with exactly K ones and study this index matrix, which now is of smaller size.

The algorithm terminates after at most $M + N - 2K$ steps. After running the algorithm we determine if the index set is minimal and can be constructed by a series of Henneberg 1 extensions. But we also obtain the pattern of extensions. Thus we obtain an efficient method of calculating the unique solution (U, V) from a data matrix X so that $W \odot (X - UV) = 0$.

3.3 Glues

Assume that the solutions to two sub-problem $\{W_1, X_1\}$ and $\{W_2, X_2\}$ are given by $\{U_1, V_1\}$ and $\{U_2, V_2\}$ respectively. To construct the solution to $\{W, X\}$, the idea is to find a transformation matrix $H \in \mathbb{R}^{K \times K}$ to transform the subspace U_2 to the same coordinate framework as the subspace U_1. Using this transformation we have

$$U_2 V_2 = (U_2 H)(H^{-1} V_2). \tag{2.40}$$

Now $U_2 H$ and $H^{-1} V_2$ are in the same coordinate framework as U_1 and V_1 respectively. The remaining problem is to solve for H. We have the following constraint, that states that U_1 and $U_2 H$ should coincide for the overlapping rows as

$$U_1(\mathtt{I}_{12}, :) = U_2(\mathtt{I}_{12}, :)H, \tag{2.41}$$

where $\mathtt{I}_1$ and $\mathtt{I}_2$ denotes the indices of overlapping rows in U_1 and U_2 respectively and $U(\mathtt{I}, :)$ denotes the sub-matrix of U by taking the rows given by $\mathtt{I}$. Similarly we have the overlapping constraints for V_1 and $H^{-1} V_2$ as

$$HV_1(:, \mathtt{J}_{12}) = V_2(:, \mathtt{J}_{12}), \tag{2.42}$$

where $\mathtt{J}_1$ and $\mathtt{J}_2$ denotes the indices of overlapping columns in V_1 and V_2 respectively.

If we have enough constraints from (2.41) and (2.42), H can be solved linearly. Two examples are if there are at least K overlapping rows or K overlapping columns. For the cases where the overlap doesn't give sufficiently many constraints, we need some extra constraint outside W_1 and W_2 to solve for the transformation matrix H.

4 Building Blocks

In this section we will describe the basic components that are used in our matrix factorization method. Our full system will be described Section 5.

4.1 Initializing Solutions

We use RANSAC to find small initial solutions. We choose a sub-problem to problem 2.38, by choosing a sub-matrix W_i of W. We further randomly select a minimal number of measurements, represented by $W_m < W_i$. These index-matrices have corresponding measurement matrices X_i and X_m. Even though we have chosen a minimal subset of measurements, this need not represent a well posed minimal problem. In order to check this, we perform Henneberg reductions as described in Section 3.2. If W_m indeed represents a minimal problem, we can, if we have a solver for this case, solve the corresponding matrix factorization of X_m. This gives a minimal solution (U_m, V_m). We can now look at how well this solution matches the other measurements in W_i by looking at the residuals $(U_m V_m - X_i) \odot W_i$. Repeating this process gives a set of initial good solutions.

4.2 Extending Solutions

If we have a solution, represented by (U_i, V_i) we can minimally extend this solution row- or columnwise using Henneberg-1 extensions, for every column (or row) that has at least K measurements. For every such column a, we randomly select K rows that are represented in the corresponding index sub-matrix W_i, and use the Henneberg-1 extension to find the new column v_a so that $\bar{V}_i = \begin{bmatrix} V_i & v_a \end{bmatrix}$. To handle outliers we check how many of the measurements that fit this new $\bar{U}_i$. If we have a substantial enough number of inliers we keep this solution, otherwise we repeat the process a number of times.

4.3 Glueing Solutions

If we have two solutions, represented by (U_i, V_i) and (U_j, V_j) we can, depending on the overlap of the two solutions, glue these solutions into one using the methods described in Section 3. Basically this can be done if there is enough information to estimate the $K \times K$ transformation matrix H so that $U_i H$ and $H^{-1} V_i$ are given in the same coordinate frame as U_j and V_j. Using a randomly selected minimal set of measurements to estimate H gives a new solution (U_k, V_k) where U_k is the union of $U_i H$ and U_j, and V_k is the union of $H^{-1} V_i$ and V_j. To handle outliers we check how many inliers we get for this new solution. Again,

if we have a substantial enough number of inliers we keep this solution, otherwise we repeat the process a number of times.

4.4 Refining Solutions

Since we use minimal solvers, and extend these solutions iteratively, we need to refine our solutions in order to avoid error propagation. We non-linearly refine our solutions, by minimizing (2.38) iteratively using Gauss Newton descent. We handle the truncation, at each step, by only optimizing over the inlier set, and then updating the inlier set using the new estimate of U and V. Since the error on the inlier set is quadratic in U and V the derivatives with respect to U and V are easily obtained.

5 Sampling Scheme

Using the building blocks from the previous section, we can now describe our full sampling scheme. The basic idea behind our method is that we will have several solutions competing against each other. These solutions will expand and merge, but at all steps we will try to be robust against outliers and inlier errors, so that we do not propagate errors. We do this by random sampling at all instances.

We assume that we have four functions available: INIT, EXTEND, GLUE and REFINE, that do initialization, extensions, glues and non-linear refinement respectively as described in Sections 4.1-4.4. We start by initializing a number of seed solutions. For each solution i, we have U_i and V_i. We then, for each of these seed solutions, attempt to extend it and refine it. If two solutions overlap, we try to glue them together. We repeat this procedure until we have at least one solution that covers the whole data matrix X. Sometimes the errors of a solution will grow during this process, and we remove solutions that have a residual norm larger than some fixed threshold. This means that we could end up with an empty solution set. In this case we re-initialize a number of seed solutions. The steps of our procedure are summarized in Algorithm 4. There is of course no guarantee that Algorithm 4 will converge, but given a well posed problem and relevant parameter settings, it is our experience that it will. We terminate after a fixed number of iterations or when we have at least one solution that covers the whole measurement matrix. We will in the experimental section validate this, on both synthetic and real data. There are a number of parameters that need to be set in order for the algorithm to work properly. The most important one is the error bound, *i.e.* the reprojection error that differentiates between an inlier and an outlier. In order to increase the robustness, we have introduced an absolute threshold on the number of inliers for the EXTEND, and GLUE functions. For each row and column the number of inliers should exceed this threshold. We have used the rank plus a small integer as threshold. We

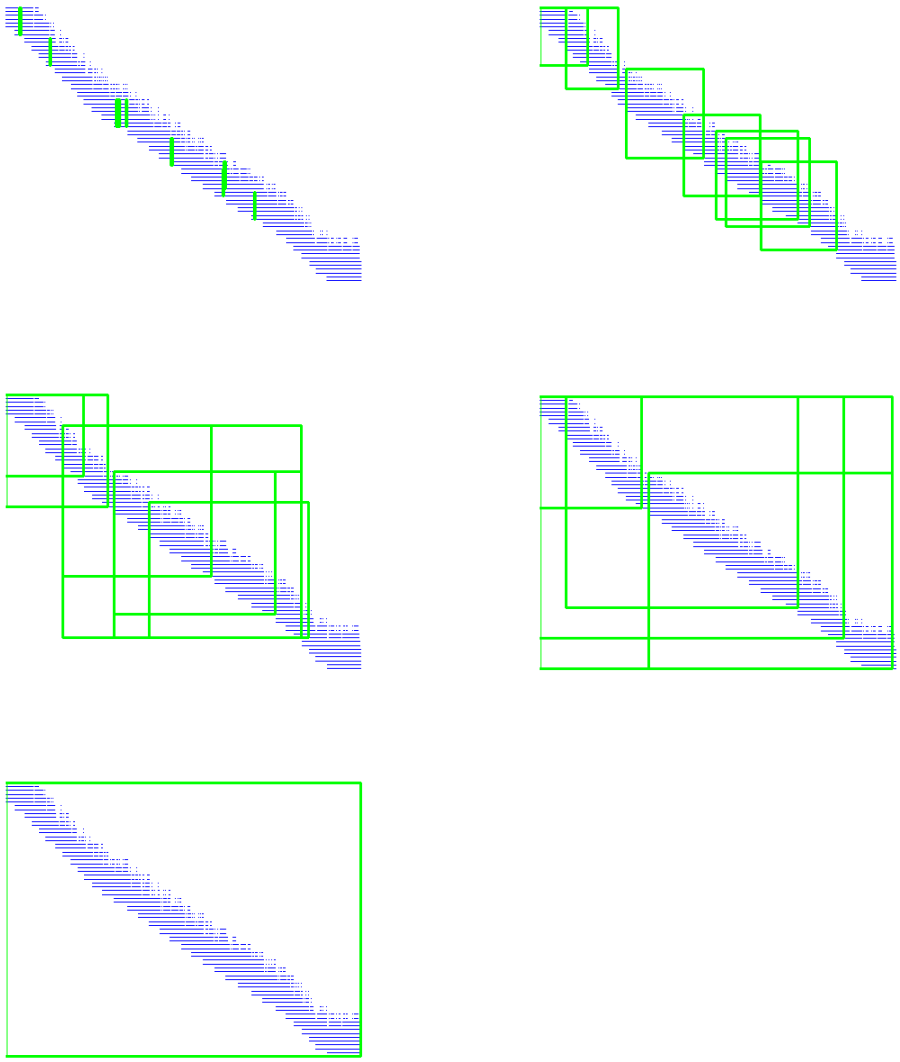


Figure 2.10: Results from running Algorithm 4 on the Dinosaur experiment (See Section 6.1 for details). Measured data points in blue, and the extent of the solutions are depicted as green boxes. The figure shows from left to right how the solution set evolves.

also need to set the number of RANSAC iterations for each of these functions, but this will mainly affect the total running time. If we assume structured data, it makes sense to smooth the index matrix W with a two-dimensional Gaussian, to obtain W_{sm}. We then sample initialization matrices guided by W_{sm} as a probability measure. Sampled positions are removed from W_{sm} to avoid multiple initial points. In [23] they do a manual block

Algorithm 4 Matrix factorization sampling scheme

Given an $M \times N$ data matrix X with index matrix W,
initialize a solution set S = INIT,
$(S = \{S_1, S_2, \ldots, S_n\}, S_i = (U_i, V_i))$.
while no S_i is of size $M \times N$ **do**
    for all i **do**
        S_i = EXTEND S_i (column-wise)
        S_i = EXTEND S_i (row-wise)
        S_i = REFINE S_i
        if $\|(U_i V_i - X_i) \odot W_i\|_F > C$ **then**
            $S = S \setminus S_i$
        end if
    end for
    for all overlapping S_i, S_j **do**
        S_k = GLUE S_i, S_j
        $S = S \cup S_k$
    end for
    if $S = \emptyset$ **then**
        Initialize n new seed solutions S = INIT.
    end if
end while

subdivision (with relatively large blocks *e.g.* 50×50). This means that the model that is fitted is very large (*e.g.* $50K + 50K - K^2$ parameters for a rank K problem) algorithm. In our approach there is no need for any explicit block subdivision, and the initial models that are fitted are much smaller. This gives a much more tractable and robust algorithm.

6 Experiments

We have applied our method to a number of different matrix factorization problems, in order to show the usefulness of it in terms of robustness, accuracy and speed.

We have compared our results to a number of state-of-the-art methods for low rank matrix factorization, namely the method of Larsson *et al.* , [21], the method of Jiang *et al.* , [23], the Truncated Nuclear Norm Regularization (TNNR-ADMM) [33], OptSpace [34], the damped Wiberg algorithm using the implementation of Okatani *et al.* [7], and the l^1 Wiberg algorithm, [11], using the C++ implementation from [35]. In the results we have included the relevant comparisons, in order to make the tables and graphs more readable. It has been previously reported (in [23]) that the methods of [33] and [34] perform much worse

for structured data patterns with high amounts of missing data. This is also our experience end hence we have omitted these results. Most of the methods do not handle outliers, and in the absence of outliers the Wiberg algorithm gives best accuracy. For this reason we mainly focus on comparison with the l^1- and l^2-versions of the Wiberg algorithm in the synthetic experiments in Section 6.3. All tests were conducted on a desktop computer running Ubuntu, with an Intel Core i7 3.6 GHz processor. The Matlab implementation of our method is available at `https://github.com/hamburgerlady/miss-ranko`.

6.1 Affine Structure from Motion

Given a full data matrix and an affine camera model it is possible to solve the structure from motion problem using factorization. The standard way of doing this is by first removing the centroid of each point in the images. This leads to a rank 3 factorization problem. When there is missing data, we can neither use SVD for factorization nor remove the centroid since this is not known. We can still write the problem as a rank 4 matrix factorization problem with missing data, see *e.g.* [36]. An affine camera is of the form

$$P_i = \begin{bmatrix} A_i \\ B_i \\ 0\ 0\ 0\ 1 \end{bmatrix}. \tag{2.43}$$

If we collect the N (homogeneous) 3D points in the $4 \times N$ matrix V, one usually writes the rank 4 problem as $UV = X$, where U is the $2M \times 4$ matrix containing the stacked camera rows A_i and B_i, and the $2M \times N$ matrix X contains the x and y coordinates of image points. However, solving this rank 4 matrix factorization problem ignores the fact that the last row of the camera matrices should be equal to $[0\ 0\ 0\ 1]$, and that the last coordinate of each homogeneous point should be equal to one. In our model we include this constraint by simply adding a row in the measurement matrix with just ones. We have found this to be a very powerful constraint, since we know it should hold even for missing data, and we use it throughout our pipeline. If the constraint is fulfilled, we can simply upgrade to the affine model by $U^a = UH$ so that the last row of U^a is equal to $[0\ 0\ 0\ 1]$. We have run Algorithm 4 on the well known Dinosaur sequence. This is a sequence that contains very little outliers, but a large amount of missing data (88%). Even though the underlying structure from motion can be (and has been) solved using a multitude of methods, the Dinosaur sequence works well as a benchmark problem for matrix factorization. We have compared our results with to those of Larsson *et al.* , [21]. In their work they also did experiments on using the nuclear norm and we have included these results here. In Jiang *et al.* , [23] it was reported that the Truncated Nuclear Norm Regularization (TNNR-ADMM) [33] and OptSpace [34] failed to recover the 2D tracks from the Dinosaur sequence. This is also our experience, as we consistently failed to recover a reasonable solution using these methods. In addition we have also run the damped Wiberg algorithm using the implementation of Okatani *et*

Table 2.2: Comparison of results from the Dinosaur and Linnaeus reconstruction experiments. The table shows the Frobenius errors using the proposed method compared to the Nuclear norm minimization, the method of Larsson *et al.* [21], and the damped Wiberg algorithm [7].

	Algorithm			
Dataset	[23]	[21]	[7]	*Proposed*
Dino ∼300 pts	99	73.3	28.0	**17.0**
Dino ∼2000 pts	-	-	144.1	**48.45**
Linnaeus ∼2000 pts	-	-	580000	**380.5**
Linnaeus ∼4000 pts	-	-	-	**543.4**

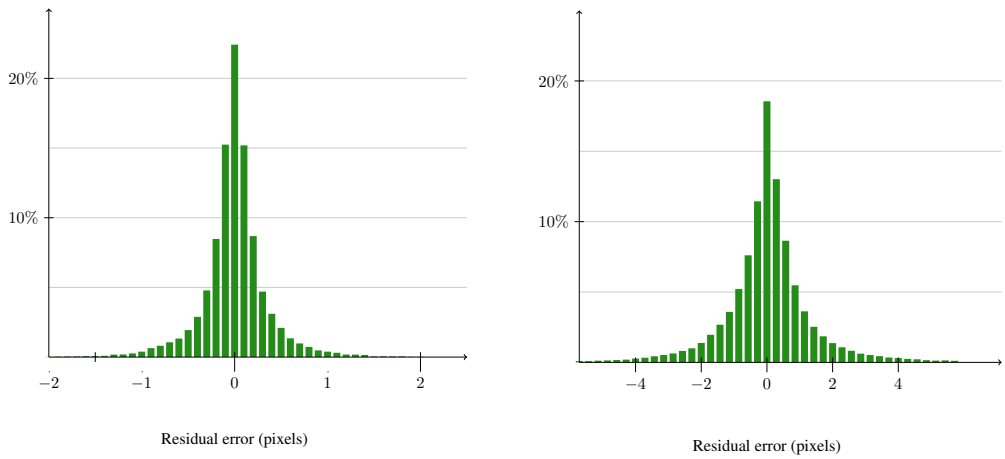


Figure 2.11: Left: Histogram of the final (non-truncated) residuals from the Dinosaur experiment using our proposed method. The results are from the larger dataset, with approximately 2000 3D points. Right: Histogram of the final (non-truncated) residuals from the Linnaeus experiment using our proposed method.

al. [7]. In Table 2.2 the Frobenius norm of the final factorizations are shown. The average running time for the Wiberg algorithm was 144 seconds, compared to around 3 seconds for our method. We have also run our method on a larger point set. For this set we didn't have access to results from [21]. We have run our method multiple times, and we always end up in the same optimum. Here the running time for the Wiberg algorithm was 4087 seconds, compared to 5.4 seconds for our method. A histogram of the residuals (in pixels) from our reconstruction is shown in Figure 2.11. Using the calibration of the projective camera model, we can upgrade our affine reconstruction to an orthographic. The resulting calibrated affine reconstruction is shown to the right in Figure 2.12. In a second experiment, we recorded a sequence of a statue of Carl Linnaeus. We extracted Harris corner points, and tracked these using the *Kanade-Lucas-Tomasi* (KLT) tracker [37]. This resulted in a sequence with

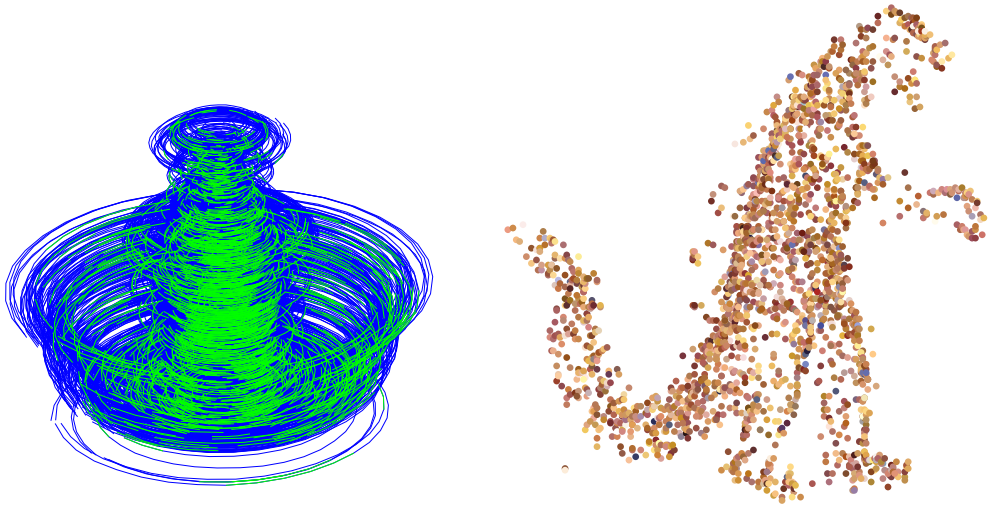


Figure 2.12: Results from the Dinosaur experiment, with approximately 2000 points. Left: Measured tracks (green) and reconstructed tracks (blue). Right: The calibrated affine reconstruction.

Figure 2.13: The figure shows three frames of the Linnaeus sequence. Also shown are the tracked Harris points (orange circles) and the reprojected points (white dots) using the proposed method. For visibility a random subset of 250 points are shown

86 images and a total of 3977 3D points. This sequence also contains closed loops, *i.e.* points were tracked from the last frames to the first. Three frames of the sequence can be seen in

Table 2.3: Comparison of the result on linear shape basis estimation using the *Book* and *Hand* dataset. The second experiment contains an additional 10% missing data. The table shows the Frobenius errors using the proposed method compared to the methods of [21], [23] and the damped Wiberg algorithm [7].

	Algorithm			
Dataset	[21]	[23]	[7]	*Proposed*
Book	0.3522	**0.1740**	**0.1740**	**0.1740**
Hand	0.8613	0.6891	**0.2802**	**0.2802**
Book-10%	8.0436	0.1772	**0.1534**	**0.1534**
Hand-10%	1.5495	0.7297	**0.2634**	**0.2634**

Figure 2.13. The results of running our algorithm on this sequence can be seen in Table 2.2. The running time for the full dataset was 39.8 s. The extracted points (orange circles) and the reprojected points (white dots) can be seen in Figure 2.13. The sequence constains approximately 1% outliers. As for the Dinosaur sequence, the Truncated Nuclear Norm Regularization (TNNR-ADMM) [33] and OptSpace [34] failed to recover reasonable 2D tracks from the Linnaeus sequence. We were not able to run the damped Wiberg algorithm on the full dataset, but the results for a subset of around half the points can be seen in Table 2.2. The running time for our method on the smaller Linnaeus sequence was 10.4 s compared to 1068 s for the Wiberg algorithm. We didn't have access to results from [21].

6.2 Linear Shape Basis Estimation

If we have non-rigid structures in a scene, a linear shape basis can be used to model the deformations. The underlying assumption is that the non-rigid deformation of the object can be represented as a linear combination of a set of shapes. Typically the size of the shape basis is much smaller than either the number of frames, or the tracked points, so the measurement matrix containing the point tracks can be factorized into a coefficient matrix and a shape basis matrix.

For our experiments we used the datasets from [21], *Book* and *Hand*. In these experiments the image points are tracked using a standard *Kanade-Lucas-Tomasi* (KLT) tracker [37]. Due to occlusions, the tracker fails after a number of frames for a subset of points, which leads to missing data. To compare with the results in [23] we use the same setup as they do, using a subset of 42 frames with 60 tracked points from the *Book* and 38 frames with 203 points from the *Hand* dataset. We then find rank-3 and rank-5 factorizations of the two datasets respectively. We ran our algorithm, and also the Wiberg algorithm using the implementation in [7]. The results can be seen in Table 2.3. The Wiberg algorithm was initialized randomly. Our method and the Wiberg minimization achieve the same optima, which are slightly better than the other methods. The reason is probably that the Wiberg and our method finds the same optimum– that we believe is the global one – since the

set is practically outlier free. The other methods do not in this case find an equally good optimum. For these smaller problems the Wiberg algorithm works well, but for larger problems it becomes intractable in terms of running time, as described in Section 6.1.

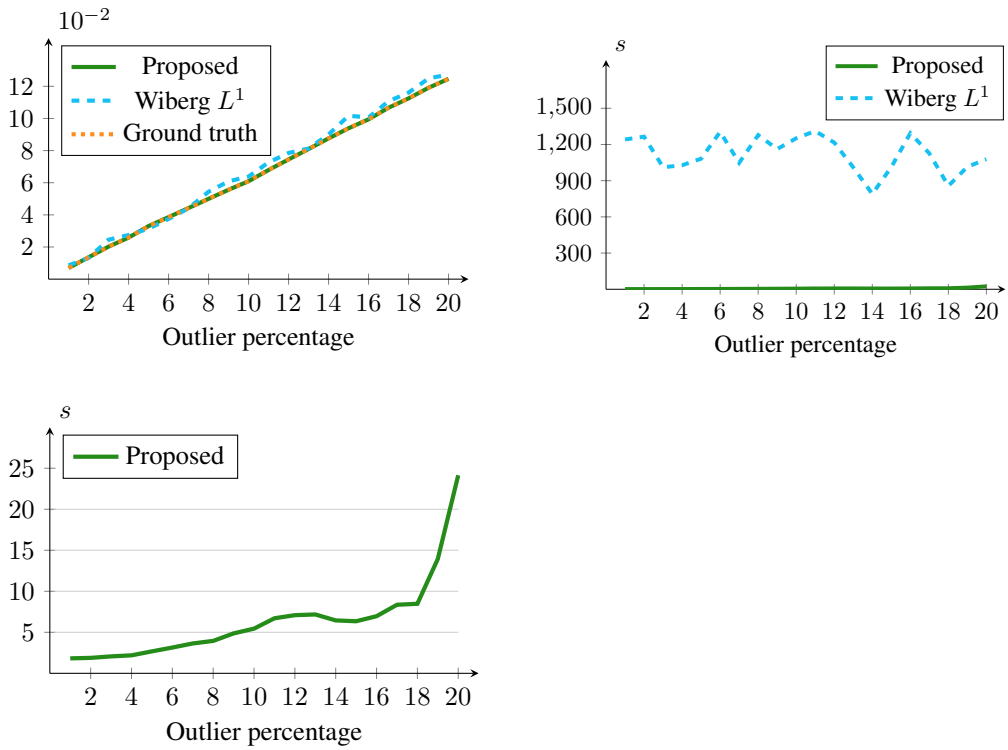


Figure 2.14: Results from the outlier test, as described in Section 6.3. Left: The graph shows the mean l^1 error for the proposed method compared to the l^1 Wiberg algorithm, [11, 35] and the oracle ground truth solution, as functions of the outlier rate. Middle: The average running times, as functions of the outlier rate, for the proposed method compared to the l^1 Wiberg algorithm, [11, 35]. Right: A magnification of the timing results for the proposed method.

6.3 Performance Tests

We have conducted a number of synthetic tests to test the performance of our method. The basic setup was done by randomly drawing two matrices $U_{M\times K}$ and $V_{K\times N}$. The product $X_0 = UV$ was then perturbed with Gaussian noise, *i.e.* $X = X_0 + \epsilon_{ij}$ with $\epsilon_{ij} \in N(0, \sigma)$. A band diagonal matrix $W_{M\times N}$ was used to prescribe which measurements were available. The bandwidth b for the matrix W (with entries w_{ij}) is defined using $w_{ij} = 0$ for $j < i - b$ or $j > i + b$. Finally, a certain percentage of the seen measurements were replaced with entries drawn randomly from a uniform distribution, to simulate gross outliers. In a

first experiment we tested the sensitivity to the proportion of outliers in the measurement matrix. We used a 300×300 measurement matrix X, with bandwidth 20. The entries were approximately between minus one and plus one, with Gaussian noise with standard deviation $1e-3$. A certain percentage of the measurements were then replaced with gross outliers. We ran Algorithm 4 and compared the results with the l^1-Wiberg algorithm, [11], using the C++ implementation from [35]. The results can be seen to the left, in Figure 2.14. We have also constructed an oracle ground truth solution to compare the results. This solution was attained by running the non-linear refinement, using the ground truth inlier set, and the ground truth U and V as starting solution. This will in general give a better optimum than U and V. One can see that both the tested algorithms perform on par. The average running times are depicted in Figure 2.14. The middle graph shows both algorithms in the same plot, and the right hand plot shows a magnification of the running times for the proposed algorithm. The l^1-Wiberg algorithm has very unattractive time complexity in terms of the size of the input measurement matrix, and we failed to run it for larger sizes than 200×200. Our method works well for moderate amounts of outliers, but as the outlier percentage increases, the RANSAC initialization will take longer and longer time, and for this test the break-down point for our method was around 20%. In a second experiment we used a similar setup, but instead of varying the outlier rate, we varied the size of the input measurement matrix X. Here we used a fixed outlier ratio of 5%. In Figure 2.15 the results can be seen. We show two versions of our algorithm, with and without using the GLUE step. The left of the figure shows the truncated l^2-error for the two versions, compared with

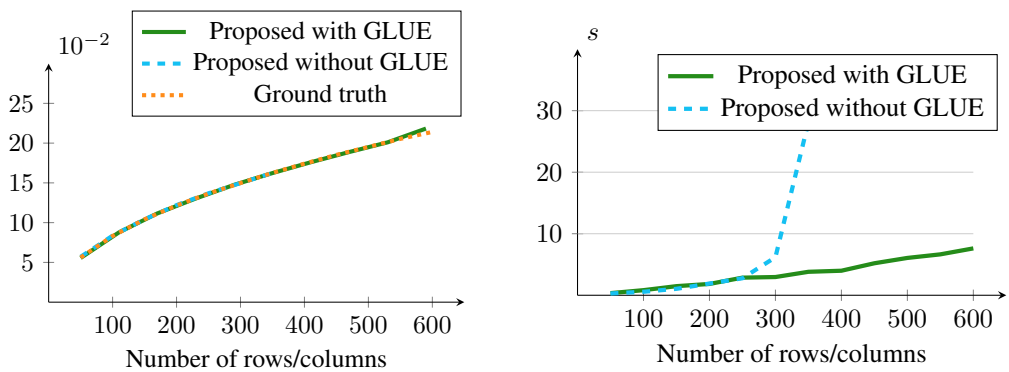


Figure 2.15: Timing results from the size test, as described in Section 6.3. The left graph shows the truncated l^2-error, as functions of the image size, for the proposed method (with and without using GLUE) compared to the oracle ground truth optimum. The right graph shows the average running times for the proposed method, with and without using GLUE.

the oracle ground truth solution, obtained in the same way as in the previous experiment. The right shows the average running times, as a function of the number of rows (equal to the number of columns) in the measurement matrix. One can see that for smaller problems,

there is no need for the GLUE. For larger problems using the GLUE leads to (in this case) near linear time complexity. As can be seen from the error plot, using the GLUE method doesn't lead to any error accumulation. We have also investigated our algorithm's

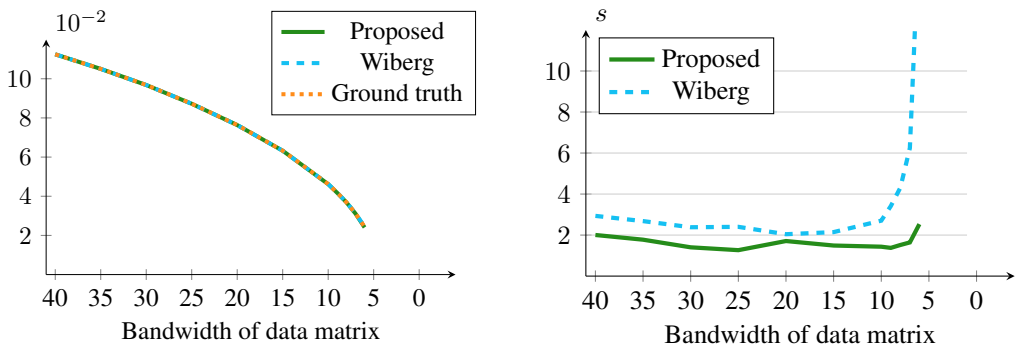


Figure 2.16: Timing results from the missing data test, as described in Section 6.3. The left graph shows the l^2-error, as functions of the bandwidth of the data, for the proposed method compared to the damped Wiberg algorithm and oracle ground truth optimum. The right graph shows the average running times for the proposed method, compared to the damped Wiberg algorithm.

dependence on the bandwidth of the measurement matrix. Using a similar setup as in the previous experiments, we constructed measurement matrices with varying bandwidth. We did not include any outliers in the data, and compared our results with the damped Wiberg algorithm. The results can be seen in Figure 2.16. Both algorithms give final norms very close to the oracle solution, but for smaller bandwidths the Wiberg algorithm has worse convergence. For very small bandwidths our algorithm becomes unstable, and the RANSAC loops will take excessive amounts of time. In this case, the breakdown point for our algorithm was for a bandwidth around 5. This will depend on the rank of the solution, which in this case was 4.

Algorithm 4 is highly parallelizable. We have conducted a simple timing experiments using our Matlab implementation. We simply change the `for`-loop to Matlabs `parfor`-loop. This runs the extensions and refinement of the initial solutions in parallel, but not the initialization. The initialization could of course also be run in parallel, but for most of our conducted experiments the initialization takes a smaller fraction of the total running time. The average running times, using different number of parallel cores, are shown in Table 2.4. We get slightly less than linear speed-ups.

Table 2.4: Timings for the linear shape basis estimation using the *Book* and *Hand* dataset. The results are based on running *parfor* in Matlab with different number of cores on an Intel Core i7 3.6 GHz processor.

	Number of cores			
Dataset	1	2	3	4
Book	0.793s	0.470s	0.382s	0.369s
Hand	28.6s	16.1s	12.2s	10.3s
Dino ∼2000 points	5.38s	3.61s	3.02s	2.19s

7 Conclusion

We have in this paper presented a method for doing robust low rank matrix factorization for structured data patterns. It can handle very large amounts of missing data, and moderate amounts of outliers. It gives results on par, or better than, using the l^1-Wiberg algorithm or the damped l^2-Wiberg algorithm, with substantial speed-up. The presented method is also trivially parallelizable. Future work includes investigating two interesting properties of our method, that we have not exploited in any detail in this paper. Firstly we have the ability to solve for multiple models in the data. For instance, if we have two rigid motions in a structure from motion sequence, we could – at least in theory – run our algorithm and find the two solutions corresponding to the two motions. Another property, that easily can be incorporated in our framework, is the ability to add additional constraints on the solution space, *e.g.* for the affine structure from motion setting we can constrain the pair-wise rows of U to be orthogonal to model a scaled orthographic camera.

Bibliography

[1] R. Basri, D. Jacobs, and I. Kemelmacher, "Photometric stereo with general, unknown lighting," *International Journal of Computer Vision*, vol. 72, no. 3, pp. 239–257, 2007.

[2] C. Bregler, A. Hertzmann, and H. Biermann, "Recovering non-rigid 3d shape from image streams," in *Conf. Computer Vision and Pattern Recognition*, 2000.

[3] R. Garg, A. Roussos, and L. Agapito, "Dense variational reconstruction of non-rigid surfaces from monocular video," in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pp. 1272–1279, IEEE, 2013.

[4] J. Yan and M. Pollefeys, "A factorization-based approach for articulated nonrigid shape, motion and kinematic chain recovery from video," *Trans. Pattern Analysis and Machine Intelligence*, 2008.

[5] C. Eckart and G. Young, "The approximation of one matrix by another of lower rank," *Psychometrika*, vol. 1, no. 3, pp. 211–218, 1936.

[6] T. Wiberg, "Computation of principal components when data are missing," in *Proc. Second Symp. Computational Statistics*, 1976.

[7] T. Okatani, T. Yoshida, and K. Deguchi, "Efficient algorithm for low-rank matrix factorization with missing components and performance comparison of latest algorithms," in *Int. Conf. Computer Vision*, pp. 842–849, 2011.

[8] A. M. Buchanan and A. W. Fitzgibbon, "Damped newton algorithms for matrix factorization with missing data," in *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 2 - Volume 02*, CVPR '05, (Washington, DC, USA), pp. 316–322, IEEE Computer Society, 2005.

[9] H. Aanaes, R. Fisker, K. Åström, and J. Carstensen, "Robust factorization," *Trans. Pattern Analysis and Machine Intelligence*, 2002.

[10] Q. Ke and T. Kanade, "Robust L_1 norm factorization in the presence of outliers and missing data by alternative convex programming," in *Conf. Computer Vision and Pattern Recognition*, 2005.

[11] A. Eriksson and A. Van Den Hengel, "Efficient computation of robust low-rank matrix approximations in the presence of missing data using the l 1 norm," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pp. 771–778, IEEE, 2010.

[12] A. D. Bue, J. M. F. Xavier, L. de Agapito, and M. Paladini, "Bilinear modeling via augmented lagrange multipliers (balm)," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 8, pp. 1496–1508, 2012.

[13] Y. Zheng, G. Liu, S. Sugimoto, S. Yan, and M. Okutomi, "Practical low-rank matrix approximation under robust L_1-norm," in *Conf. Computer Vision and Pattern Recognition*, 2012.

[14] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *CoRR*, vol. abs/0805.4471, 2008.

[15] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?," *Journal of ACM*, 2009.

[16] Z. Lin, M. Chen, L. Wu, and Y. Ma, "The augmented lagrange multiplier method for exact recovery of a corrupted low-rank matrices," *CoRR*, vol. abs/1009.5055, 2009.

[17] C. Olsson and M. Oskarsson, "A convex approach to low rank matrix approximation with missing data," in *Scandinavian Conf. on Image Analysis*, 2011.

[18] R. Cabral, F. D. la Torre, J. P. Costeira, and A. Bernardino, "Unifying nuclear norm and bilinear factorization approaches for low-rank matrix decomposition," in *Int. Conf. Computer Vision*, 2013.

[19] L. W. Mackey, M. I. Jordan, and A. Talwalkar, "Divide-and-conquer matrix factorization," in *Advances in Neural Information Processing Systems*, pp. 1134–1142, 2011.

[20] C. Julià, A. D. Sappa, F. Lumbreras, J. Serrat, and A. López, "An iterative multiresolution scheme for sfm with missing data," *Journal of Mathematical Imaging and Vision*, vol. 34, no. 3, pp. 240–258, 2009.

[21] V. Larsson, C. Olsson, E. Bylow, and F. Kahl, "Rank minimization with structured data patterns," in *ECCV*, 2014.

[22] V. Larsson and C. Olsson, "Convex envelopes for low rank approximation," in *EMMCVPR 2015*, 2015.

[23] F. Jiang, M. Oskarsson, and K. Åström, "On the minimal problems of low-rank matrix factorization," in *Conf. Computer Vision and Pattern Recognition*, 2015.

[24] F. Kahl, A. Heyden, and L. Quan, "Projective reconstruction from minimal missing data," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, no. 4, pp. 418–424, 2001.

[25] M. Oskarsson, K. Åström, and N. C. Overgaard, "Classifying and solving minimal structure and motion problems with missing data," in *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, vol. 1, pp. 628–634, IEEE, 2001.

[26] N. Guilbert and B. A., "Batch recovery of multiple views with missing data using direct sparse solvers," in *British Machine Vision Conference*, 2003.

[27] M. Oskarsson, K. Åström, and A. Torstensson, "Prime rigid graphs and multidimensional scaling with missing data," in *Pattern Recognition (ICPR), 2014 22nd International Conference on*, pp. 750–755, Aug 2014.

[28] J. Karlsson and K. Åström, "Mdl patch correspondences on unlabeled images with occlusions," in *CVPR, Computer Vision and Pattern Recognition*, 2008.

[29] D. Nistér, F. Kahl, and H. Stewénius, "Structure from motion with missing data is *NP*-hard," in *Int. Conf. Computer Vision*, (Rio de Janeiro, Brazil), 2007.

[30] A. Blake and A. Zisserman, *Visual Reconstruction*. MIT Press, Cambridge, USA, 1987.

[31] G. Laman, "On graphs and rigidity of plane skeletal structures," *Journal of Engineering mathematics*, vol. 4, no. 4, pp. 331–340, 1970.

[32] L. Henneberg, *Die graphische Statik der starren Systeme*, vol. 31. BG Teubner, 1911.

[33] Y. Hu, D. Zhang, J. Ye, X. Li, and X. He, "Fast and accurate matrix completion via truncated nuclear norm regularization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 9, pp. 2117–2130, 2013.

[34] R. H. Keshavan, S. Oh, and A. Montanari, "Matrix completion from a few entries," *CoRR*, vol. abs/0901.3150, 2009.

[35] D. Strelow, "General and nested wiberg minimization," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pp. 1584–1591, IEEE, 2012.

[36] G. Wang, J. S. Zelek, Q. J. Wu, and R. Bajcsy, "Robust rank-4 affine factorization for structure from motion," in *Applications of Computer Vision (WACV), 2013 IEEE Workshop on*, pp. 180–185, IEEE, 2013.

[37] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proceedings of the 7th International Joint Conference on Artificial Intelligence - Volume 2*, IJCAI'81, (San Francisco, CA, USA), pp. 674–679, Morgan Kaufmann Publishers Inc., 1981.

Paper v

Towards Real-time Time-of-Arrival Self-Calibration using Ultra-Wideband Anchors

Kenneth Batstone, Magnus Oskarsson and Kalle Åström
Centre of Mathematical Sciences, Lund University, Lund, Sweden

Abstract: Indoor localisation is a currently a key issue, from robotics to the Internet of Things. With hardware advancements making Ultra-Wideband devices more accurate and low powered (potentially even passive), this unlocks the potential of having such devices in common place around factories and homes, enabling an alternative method of navigation. Therefore, anchor calibration indoors becomes a key problem in order to implement these devices efficiently and effectively.

In this paper, we study the possibility for sequentially gathering Ultra-Wideband Time-of-Arrival measurements and using previously studied robust solvers, merge solutions together in order to calculate anchor positions in 3D in real-time. Here it is assumed that there is no prior knowledge of the anchor positions. This is then validated using Ultra-Wideband Time-of-Arrival data gathered by a Bitcraze Crazyflie quadcopter in 2D motion, 3D motion and full flight.

1 Introduction

Indoor positioning and navigation is a well known problem. In the modern world, systems such as GPS are heavily relied on for the navigation and positioning of people, vehicles and robots to name a few. Unfortunately, in highly urban areas, such as New York, tall buildings interfere with the accuracy of positioning and this becomes worse once inside any building. Once inside a building it can, at times, be impossible to acquire signals from the GPS satellites. When this occurs navigation users must use an alternative system. There are currently many options to overcome this problem but they all come with their own flaws. A large focus has been using the signal strength of Wi-Fi networks to perform positioning since the infrastructure currently exists in most buildings but due to the nature of radio signals, they have a low accuracy and with distance, the errors increase exponentially, [1]. Other options include Bluetooth devices (that have a short range) and dead reckoning approaches (that decrease in accuracy over time due to perturbations in errors).

Now, a new era of wireless communications, **5G** is hoping to bridge the gap to enable reliable indoor positioning. One such technology which is commercially available is Ultra-Wideband. The devices which use Ultra-Wideband are low powered and perform 2-way timing in order to obtain high precision in positioning, [2], between 2 devices. With further hardware advancements, Ultra-Wideband devices aim to be more accurate and lower powered (potentially even passive, [3]). This unlocks the potential of having such devices in common place around factories and homes, enabling an alternative method of navigation indoors for people and Internet-of-Things (IoT) devices.

In this paper we present new research on methods for time-of-arrival (TOA) self-calibration problem. Here we exploit the factorisation of the receiver-transmitter distance matrix in order to enable real-time anchor calibration. This method is then applied to measurements taken using a Bitcraze, Crazyflie quadcopter with Ultra-Wideband anchors in order to determine if the proposed method is feasible in a real world situation, shown in Figure 2.24. These devices use the Decawave DWM1000 chip. For our experiments, the Ultra-Wideband anchors have the notation as either receiver or anchor, and the Crazyflie quadcopter as the transmitter.

The TOA self-calibration problem, is the problem of determining the positions of a number of receivers and transmitters given only receiver-transmitter distances. Here, there is no prior knowledge of the anchor positions. Hence, it is closely related to the *anchor-free* sensor network localisation problem [4] but differs, since our problem corresponds to a bipartite graph network, where [4] the network structure is a general graph. In [5], the authors explore the problem of embedded minimal rigid graphs in a larger general graph in order to solve the larger graph in the presence of missing information, which is applicable to the TOA problem. The TOA problem also has certain similarities with the problem of

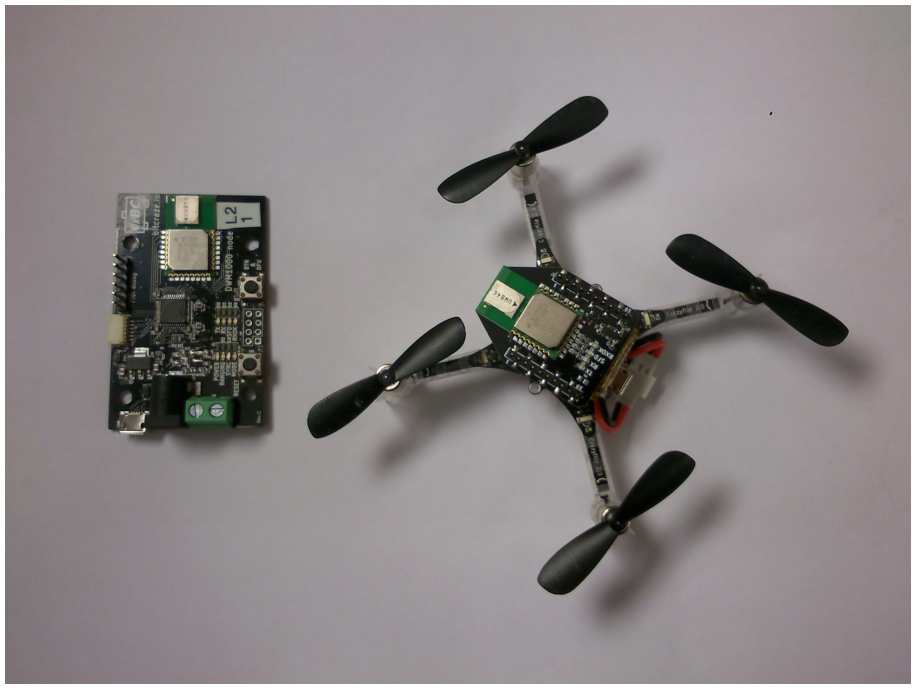


Figure 2.17: Image of the Ultra-Wideband anchor and Bitcraze Crazyflie quadcopter respectively from left to right.

determining a set of points given all inter-point distances, which is usually solved using multi-dimensional scaling [6].

Anchor-free sensor network calibration with time-of-arrival measurements has been investigated in a number of studies. Graph rigidity was explored in [4] to find a fold-free graph embedding, and the solution was then refined using mass-spring based optimisation. In [7], a semi-definite programming formulation and solution was proposed for TOA measurements, with or without anchors. Both of these methods are general for any solvable network structure, [8]. In an other line of work, there has been focused on sensor networks with a bipartite structure. For this type of network structure, one also aims to calculate the minimal problem, i.e. minimal number of receivers and transmitters required for the problem to be solvable. Note that for this problem, the roles of receivers and transmitters are equivalent. The minimal cases were studied in [9], which found that the minimal case of a 2D network is three transmitters and three receivers. The minimal problems for the 3D case are given in [10]. Here the minimal configuration of receivers and transmitters are $(4, 6)$, $(5, 5)$ and $(6, 4)$ respectively. There are in general 38, 42 and 38 solutions respectively for the three types of problems. However, no practical methods for general 3D positions are given. There are a few results on algorithms for actually determining the positions from distances, most notably [11, 12]. In [13], a non-minimal linear solution to the 3D TOA

self-calibration problem is derived for 10 (4) receivers and 4 (10) transmitters. In [14, 15] a solution is given to the TOA self-calibration problem, if one may additionally assume that one of the receivers coincides with the position of one of the transmitters.

Studying these minimal cases is of theoretical importance and further more essential when developing fast and stable algorithms based on robust estimation methods like RANSAC [16], in the presence of outliers in the measurements. As will be shown in the following sections, one important part of our system exploits that the so-called *compaction matrix* should have a certain rank. Low rank matrix factorisation has a long standing history. Truncating the singular value decomposition of the measurement matrix has been shown to give the optimal solution under the l^2-norm for complete data, see [17]. The work in [18] was the first to consider missing data. Robustness to outliers has been considered in [19, 20, 21, 22]. Most methods mentioned above are based on alternating optimisation and are prone to get trapped in local minima. Recently, several works [23, 24, 25] re-formulate the problem to minimise the convex surrogate of the rank function, that is, the nuclear norm. For applications with a given fixed rank, the nuclear norm based methods usually perform inferior to the bilinear formulation-based methods [26]. A few recent works [27, 28, 29, 30] explore the idea of dividing the whole matrix into overlapping sub-blocks and combine the sub-block solutions. Minimal cases for low rank matrix factorisation, for missing data, were investigated in [31]. This paper is a continuation of our paper [32], where a RANSAC paradigm was used in conjunction with minimal solvers and explored in order to obtain a robust and fast solution of the TOA self-calibration problem.

The novel concept of this paper, is that we can use our robust method to find a solution for a short dataset, then merge this solution sequentially to a previous solution. By doing this we see a convergent nature of calibrating the anchor positions in 2D and 3D as more data is collected. This has many added advantages. One such advantage is that it can be determined when there has been enough data acquired to solve the self calibration problem. In doing so, after the problem has been solved the user could switch to computationally lighter positioning algorithms to increase the performance. It also has the advantage of being implemented on large scales. As long as the solutions are connected, the solutions from one part of a building would be in the same coordinate system as another part of the building, but only the information from one part of the building is required to form a positioning solution, hence reducing the complexities of data storage and optimisations. In addition, since we are dealing with short datasets, these are quickly optimised to their local minima, so we have found overall improvements to our anchor positions calculations.

2 Basic Geometry

We will now describe the basic underlying geometry of our problem. Let $\mathbf{r}_i$, $i = 1, \ldots, m$ and $\mathbf{s}_j$, $j = 1, \ldots, n$ be the spatial coordinates of m receivers (e.g. Ultra-Wideband anchors) and n transmitters (e.g. Crazyflie quadcopter), respectively. For measured time of arrival t_{ij} from transmitter $\mathbf{r}_i$ and receiver $\mathbf{s}_j$, we have $vt_{ij} = \|\mathbf{r}_i - \mathbf{s}_j\|_2$ where v is the speed of measured signals and $\|.\|_2$ is the ℓ^2-norm. The speed v is assumed to be known and constant. We further assume that we, at each receiver can distinguish which transmitter j each event is originating from. This can be done e.g. if the signals are temporally separated or using different frequencies. We will in the following work with the distance measurements $d_{ij} = vt_{ij}$. It is quite common that such data contains both missing data from poor signal communications and outliers due to inaccuracies of the hardware measurements. The TOA calibration problem can then be defined as follows,

Problem 4. *(Time-of-Arrival Self-Calibration) Given absolute distance measurements*

$$d_{ij} = \|\mathbf{r}_i - \mathbf{s}_j\|_2 + \epsilon_{i,j}, \tag{2.44}$$

where the receiver positions are defined as $\mathbf{r}_i$, $i = 1, \ldots, m$ *and transmitter positions as* $\mathbf{s}_j$, $j = 1, \ldots, n$. *Here the errors* $\epsilon_{i,j}$ *are assumed to be either* **inliers**, *in which case the errors are small (*$\epsilon_{i,j} \in N(0, \sigma)$*) or* **outliers**, *in which case the measurements are way off.*

Here we will use the set W_i for the indices (i, j) corresponding to the inlier measurements.

We will now show how the TOA calibration problem is solved generally. From many types of media, a transmitter-receiver distance will be acquired, d_{ij}. Since this can be assumed to be real and positive, it can be squared as follows,

$$d_{ij}^2 = (\mathbf{r}_i - \mathbf{s}_j)^T(\mathbf{r}_i - \mathbf{s}_j) = \mathbf{r}_i^T\mathbf{r}_i + \mathbf{s}_j^T\mathbf{s}_j - 2\mathbf{r}_i^T\mathbf{s}_j. \tag{2.45}$$

The problem is then reformed according to the following invertible linear combinations of d_{ij}^2:

$$\mathbf{B} = \begin{pmatrix} d_{11}^2 & d_{12}^2 - d_{11}^2 & \ldots & d_{1n}^2 - d_{11}^2 \\ d_{21}^2 - d_{11}^2 & & & \\ \ldots & & \hat{\mathbf{B}} & \\ d_{m2}^2 - d_{11}^2 & & & \end{pmatrix}, \tag{2.46}$$

where the *compaction matrix* $\hat{\mathbf{B}}$ is an $(m-1) \times (n-1)$ matrix with entries as $\hat{B}_{ij} = \dfrac{d_{i,j}^2 - d_{i1}^2 - d_{1j}^2 + d_{11}^2}{-2}$, with $i = 2, \ldots m$ and $j = 2, \ldots, n$. The other elements in $\mathbf{B}$ are used as constraints for the solution.

The factorisation can then be interpreted as follows. Let $\mathbf{R}_i = \left[(\mathbf{r}_i - \mathbf{r}_1)\right]$ and $\mathbf{S}_j = \left[(\mathbf{s}_j - \mathbf{s}_1)\right]$. Here $\hat{\mathbf{B}} = \mathbf{R}^T\mathbf{S}$ with $\mathbf{R}_i$ as columns of $\mathbf{R}$ and $\mathbf{S}_j$ as columns of $\mathbf{S}$. Since we assume that $\mathbf{R}$ and $\mathbf{S}$ are in a 3D affine space, the matrix $\hat{\mathbf{B}}$ has rank 3 at most. This also implies that in order to solve the problem, it is required that $m \geq 4$ and $n \geq 4$. By factorising $\hat{\mathbf{B}}$, we can compute the vectors to all receivers and transmitters from unknown initial/reference positions ($\mathbf{r}_1$ and $\mathbf{s}_1$).

By fixing $\mathbf{r}_1$ at the origin and $\mathbf{s}_1$ as a vector from the origin, in terms of an affine transformation matrix $\mathbf{L}$ and vector $\mathbf{b}$, the problem is reformulated as follows,

$$\begin{aligned} &\mathbf{r}_1 = 0,\ \mathbf{s}_1 = \mathbf{L}\mathbf{b}, \\ &\mathbf{r}_i = \mathbf{L}^{-T}\tilde{\mathbf{R}}_i,\ i = 2 \ldots m, \\ &\mathbf{s}_j = \mathbf{L}(\tilde{\mathbf{S}}_j + \mathbf{b}),\ j = 2 \ldots n, \end{aligned} \tag{2.47}$$

where $\tilde{\mathbf{R}} = \mathbf{L}^T\mathbf{R}$, $\tilde{\mathbf{S}} = \mathbf{L}^{-1}\mathbf{S}$, and hence $\hat{\mathbf{B}} = \tilde{\mathbf{R}}^T\mathbf{L}^{-1}\mathbf{L}\tilde{\mathbf{S}} = \mathbf{R}^T\mathbf{S}$.

Using this parametrisation, the equations from matrix $\mathbf{B}$, (2.46) become

$$\begin{aligned} d_{11}^2 &= (\mathbf{r}_1 - \mathbf{s}_1)^T(\mathbf{r}_1 - \mathbf{s}_1) = \mathbf{s}_1^T\mathbf{s}_1 = \mathbf{b}^T\mathbf{L}^T\mathbf{L}\mathbf{b} \\ &= \mathbf{b}^T\mathbf{H}^{-1}\mathbf{b}, \end{aligned} \tag{2.48}$$

$$\begin{aligned} d_{1j}^2 - d_{11}^2 &= \mathbf{s}_j^T\mathbf{s}_j - \mathbf{s}_1^T\mathbf{s}_1 = \tilde{\mathbf{S}}_j^T\mathbf{L}^T\mathbf{L}\tilde{\mathbf{S}}_j + 2\mathbf{b}^T\mathbf{L}^T\mathbf{L}\tilde{\mathbf{S}}_j \\ &= \tilde{\mathbf{S}}_j^T\mathbf{H}^{-1}\tilde{\mathbf{S}}_j + 2\mathbf{b}^T\mathbf{H}^{-1}\tilde{\mathbf{S}}_j, \end{aligned} \tag{2.49}$$

$$\begin{aligned} d_{i1}^2 - d_{11}^2 &= \mathbf{r}_i^T\mathbf{r}_i - 2\mathbf{r}_i^T\mathbf{s}_1 = \tilde{\mathbf{R}}_i^T(\mathbf{L}^T\mathbf{L})^{-1}\tilde{\mathbf{R}}_i - 2\mathbf{b}^T\tilde{\mathbf{R}}_i \\ &= \tilde{\mathbf{R}}_i^T\mathbf{H}\tilde{\mathbf{R}}_i - 2\mathbf{b}^T\tilde{\mathbf{R}}_i, \end{aligned} \tag{2.50}$$

where the symmetric matrix $\mathbf{H} = (\mathbf{L}^T\mathbf{L})^{-1}$. With this parametrisation, there are in total 9 unknowns (6 and 3 unknowns for $\mathbf{H}$ and $\mathbf{b}$, respectively), and hence a solution can be found. Since this solution has its own coordinate system, with prior knowledge such as gyroscope data, this can be transformed back to the original coordinate system.

3 Non-Linear Optimisation Approaches

In the development of the different systems for robust estimation, we use several local optimisation techniques. In particular we use methods for local optimisation of the type

$$\min_{\mathbf{r},\mathbf{s}} \sum_{(i,j)\in\tilde{W}} f(d_{i,j} - ||\mathbf{r}_i - \mathbf{s}_j||_2), \tag{2.51}$$

where $f(r)$ is chosen to be (i) $f(r) = r^2$ (l^2-norm), (ii) $f(r) = |r|$ (l^1-norm) or (iii) $f(r) = \min(r^2, T)$ (truncated l^2-norm). If the subset $\tilde{W}$ of the measurements contains no outliers and if the starting point is good, then the l^2-norm can give good estimates. Optimising using the l^1-norm is less sensitive to the subset $\tilde{W}$ containing outliers, but still requires a reasonably good starting point to converge to a good solution. Local optimisation of the truncated l^2-norm is even more sensitive to having a good starting point. Nevertheless, these local optimisation techniques are important components for designing robust systems.

4 Obtaining Initial Estimates

Finding the optimal solution to problem 5, in the presence of outliers and missing data is a highly non-convex problem. We are thus dependent on finding good initial starting solutions, for the optimisation methods from the previous section to work. We will in this section describe the different initialisation methods that we have used in our experiment. In the next section we will describe our latest contribution to the initialisation problem.

Arguably, the most straight-forward way to initialise a solution, is to simply randomly place all receivers and transmitters within some space. This usually gives poor initial estimates, and the local optimisation will be prone to get stuck in local minima. A slight improvement to this idea, is to use multiple restarts and optimise from each initial position, and then in the end choose the best solution.

Another way of initialising, that we have explored, is using the rank constraint on the compaction matrix. Here one can use many existing methods for doing the low rank matrix factorisation. One important draw-back of these methods, is that we need to have at least one row and one column of the data matrix completely known, and without outliers. The last criteria is of course hard to check. If all data is known, the optimal low rank factorisation is given by singular value decomposition (SVD) of the data matrix. A heuristic for handling missing data, is simply to fill in the missing data with some random values that follow the statistics of the other known measurements. One can then use SVD to obtain an initial estimate. This can be used directly to find the solution to the original problem as described in section 2. Alternatively, the initial low rank matrix factorisation can be refined using the Wiberg algorithm, [18].

5 Random Sampling Paradigm

The RANSAC or hypothesise and test paradigm, has proven to be useful in situations where there are outliers in the data, [16]. In this paradigm, a subset of the data is used to estimate

the unknown parameters. The remainder of the data is then used to verify or falsify the parameters. This is typically repeated a fixed number of iterations. The parameters that give the largest number of inliers are then usually used as an initial estimate for the subsequent non-linear optimisation of the parameters.

For Problem 5, there are several ways one could implement the hypothesis and test paradigm.

One idea would be to use efficient algorithms for determining receiver and transmitters positions from minimal data, [10]. Although this solver and the test is relatively fast, we propose an alternative to this approach. In one of our recent papers the idea is to find a fast way to hypothesise and test. We will use the rank constraints of the compaction matrix to do this. Our method is described in [32].

6 Merging Solutions

Assume that there are two sequential solutions $\hat{\mathbf{B}}_1 = \mathbf{R}_1^T\mathbf{S}_1$ and $\hat{\mathbf{B}}_2 = \mathbf{R}_2^T\mathbf{S}_2$. Within these two solutions there exists a subset of data that overlaps denoted by a subscript, i. Since the two solutions are in different coordinate systems, due to coordinate dis-ambiguities in our solutions, then the idea is to find a transformation matrix $Q \in \mathbb{R}^{K\times K}$ to transform $\mathbf{R}_2$ into the same coordinate system as $\mathbf{R}_1$. Using this transformation matrix, the solution can be transformed to another coordinated system where the solution is just as valid, i.e.

$$\mathbf{R}_2^T\mathbf{S}_2 = \mathbf{R}_2^T QQ^{-1}\mathbf{S}_2. \tag{2.52}$$

Since there is an overlapping region, i, in the solutions, it can be assumed that

$$\mathbf{R}_{1,i}^T\mathbf{S}_{1,i} = (\mathbf{R}_{2,i}^T Q)(Q^{-1}\mathbf{S}_{2,i}), \tag{2.53}$$

and hence,

$$\mathbf{R}_{1,i}^T = \mathbf{R}_{2,i}^T Q, \tag{2.54}$$

$$\mathbf{S}_{2,i} = Q\mathbf{S}_{1,i}. \tag{2.55}$$

Provided there is enough data in the overlapping region (2.54) and (2.55), then the transformation matrix Q can be solved linearly. The two solutions can then be merged, the overlapping region $\mathbf{R}_{1,i}, \mathbf{S}_{1,i}$ is updated using (2.54), (2.55) and the rest of $\mathbf{R}_2, \mathbf{S}_2$ are concatenated onto the previous solution using (2.52). Since $\mathbf{R}$ and $\mathbf{S}$ are linear combinations of $\mathbf{r}$ and $\mathbf{s}$, then the compaction matrix doesn't need to be calculated in order to solve this, instead $\mathbf{r}$ and $\mathbf{s}$ are used to calculate Q. The details of our algorithm are explained in Algorithm 6.

Algorithm 5 Our Solution merging scheme

1: Select Overlapping region such that it is assumed that $\mathbf{r}_1 = \mathbf{r}_2$
2: Calculate the transformation matrix in a RANSAC way to obtain the best Q.
3: **If:** when applying the best transformation matrix Q, the new solution does not fit well i.e. $||\mathbf{r}_{1,i}^T - \mathbf{r}_{2,i}^T Q||_2 + ||\mathbf{s}_{1,i}^T - Q\mathbf{s}_{2,i}^T||_2 > \epsilon$, for some sufficiently small ϵ, then $\mathbf{r}_1$ remains the same and $\mathbf{s}_2$ are trilaterated points using $\mathbf{r}_1$.
4: **Else:** Calculate the new $\mathbf{r}_{1,i}$ and $\mathbf{s}_2$ according to (2.54) and (2.55).
5: Optimise over the range of both solutions.

7 Experimental Setup

To test our method 3 experiments were conducted using a Bitcraze Crazyflie quadcopter and their Loco-positioning system which consists of Ultra-Wideband chips on the anchors and quadcopter. Six anchors were positioned around the rooms and for two of the experiments, measured using a laser distance meter to obtain ground truth positions of the anchors in the 2D experiment and full flight experiment with an error of ± 2 *mm*.

The other experiment was conducted in a MOCAP studio to record the ground truth flight-path as well as the anchor positions with an error of ± 1 *mm*.

Distance measurements from the quadcopter to all the anchors were measured at a frequency of 30 *Hz*.

Two of the experimental environments were rooms in an office block, which reduced the chances of large outliers or missing data, but the 2D experiment was performed in a large open space in an office block with meeting rooms in the centre in order to provoke large amounts of missing data.

The experiments were conducted by moving the quadcopter, by hand, around the room except for the full flight experiment, where the quadcopter was flown. The distance measurements were recorded so that they may be processed offline. Our algorithms do not require any prior knowledge of anchor or quadcopter positions. The only requirement is that our minimal solver (5,5) is satisfied for the 3D cases and (3,3) for the 2D case.

8 Experimental Evaluation

Once the measurements were taken, they were streamed into our algorithm to simulate real-time acquisition. At all optimisation steps, l^2 optimisation on the inlier set was performed. We also added a smoothness prior in the optimisation. This prior is based on minimising

acceleration, according to

$$res_a = \frac{1}{\sigma_a^2} \sum_{j=2}^{n-1} \|\mathbf{s}_{j-1} - 2\mathbf{s}_j + \mathbf{s}_{j+1}\|_2^2, \tag{2.56}$$

where σ_a is a parameter controlling the strength of the smoothness prior. This is reasonable since the motion of the quadcopter is continuous.

8.1 MOCAP Experiment

When streaming the data, a number of buffer lengths were tested, in order to see if this had an effect on the rate of convergence. The result of this can be see in Figure 2.18. To show this result a Procrustes fitting was used to find a transformation between the ground truth anchor positions and the final calculated anchor positions. This transform was then applied to all calculated solutions retrospectively so that they are in equivalent coordinate systems. In the Procrustes fitting only rotation and translation were allowed.

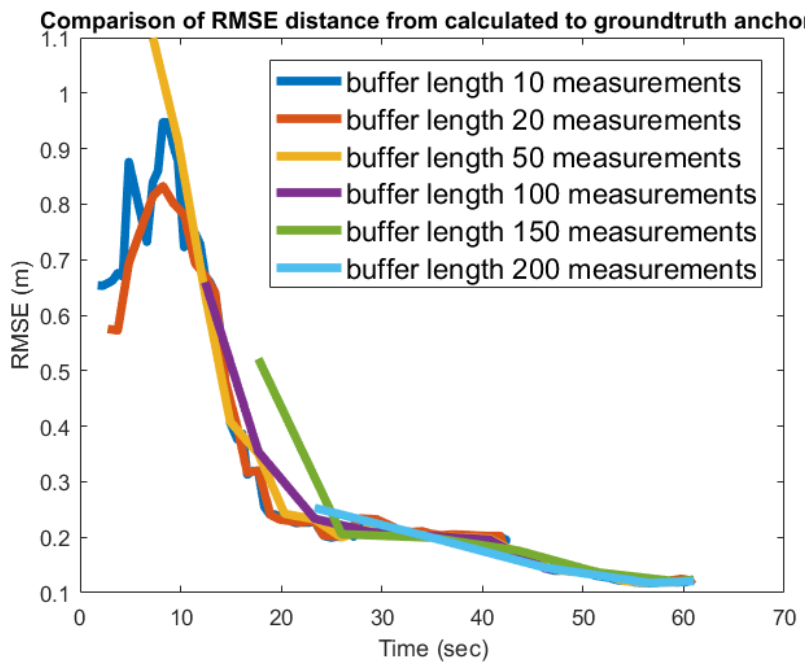


Figure 2.18: This figure illustrates the Root Mean Squared Error (RMSE) distance between the calculated to ground truth anchor positions for different buffered measurement sizes.

A buffer length of 20 measurements was then chosen for the subsequent experiments.

Using the same set-up and data from the MOCAP studio, solutions were found for each 20 measurement and merged together using 10 overlapping measurements. As before a retrospective Procrustes fitting was applied. Here the (5,5) minimal solver was used.

To show the convergence of each of the anchor positions, the distance from the calculated anchor position to the ground truth anchor positions are plotted for each step, as shown in Figure 2.19.

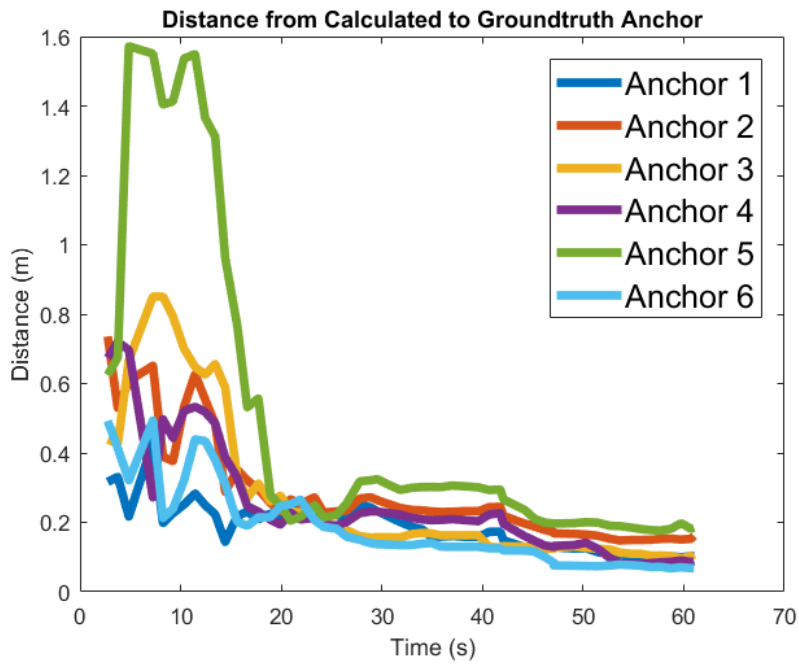


Figure 2.19: This figure illustrates the distance between the calculated to ground truth anchor positions.

Here the final distance from the calculated anchor position to the ground truth anchor positions are as follows, $(0.1026, 0.1612, 0.1078, 0.0888, 0.1845, 0.0675)m$ for anchors 1 to 6 respectively.

The final solution is shown in Figure 2.20. In this figure, three instances from sequentially calculating and merging solutions are shown with the last being the final solution. In this dataset there was 0.61% of the data missing. For the final solution, the distance from each quadcopter positions to the corresponding ground truth position was calculated. A histogram of the errors can be seen in Figure 2.21.

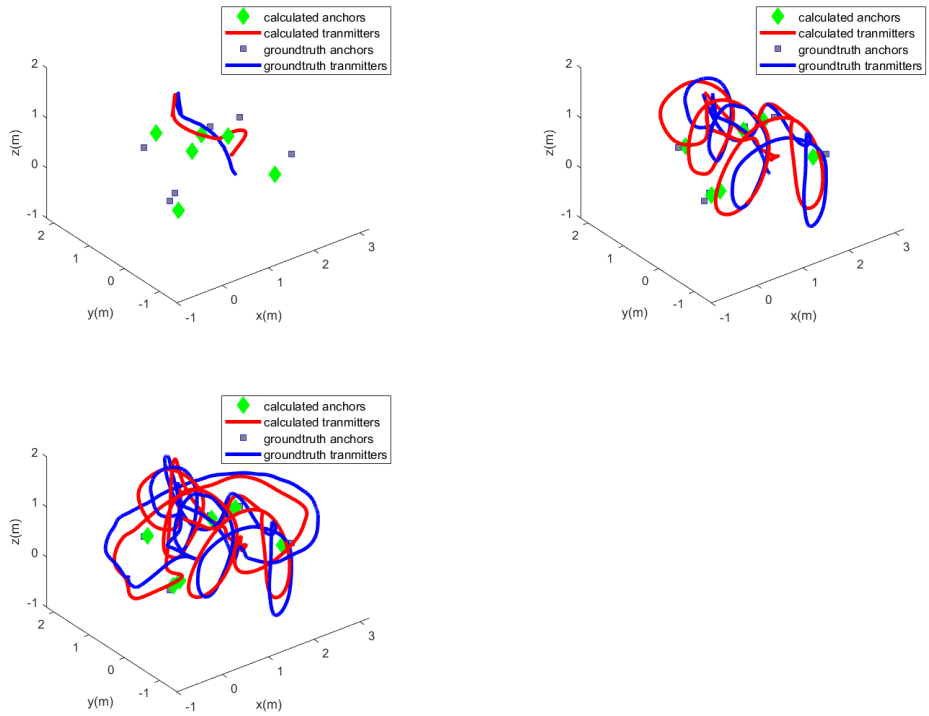


Figure 2.20: The figure illustrates the estimated anchor positions and quadcopter positions. These are overlaid on the ground truth anchor positions and flightpath. Each graph represents a different, (increasing) time instance, the bottom corresponding to the final solution.

8.2 Full Flight Experiment

During this experiment, a video recording of the experiment was made. In Figure 2.22, three of the video frames are shown with the 3D-points projected onto the image. The main purpose is to be able to visualise the convergence of the anchor positions during a flight. Here the (5,5) minimal solver was used.

Here the distance from the final calculated anchor position to the ground truth anchor positions are as follows, $(0.1439, 0.1154, 0.1501, 0.1664, 0.1477, 0.1842)m$, for anchors 1 to 6 respectively. In this dataset there was 0.01% of the data missing. Again, a retrospective Procrustes fitting was applied as before.

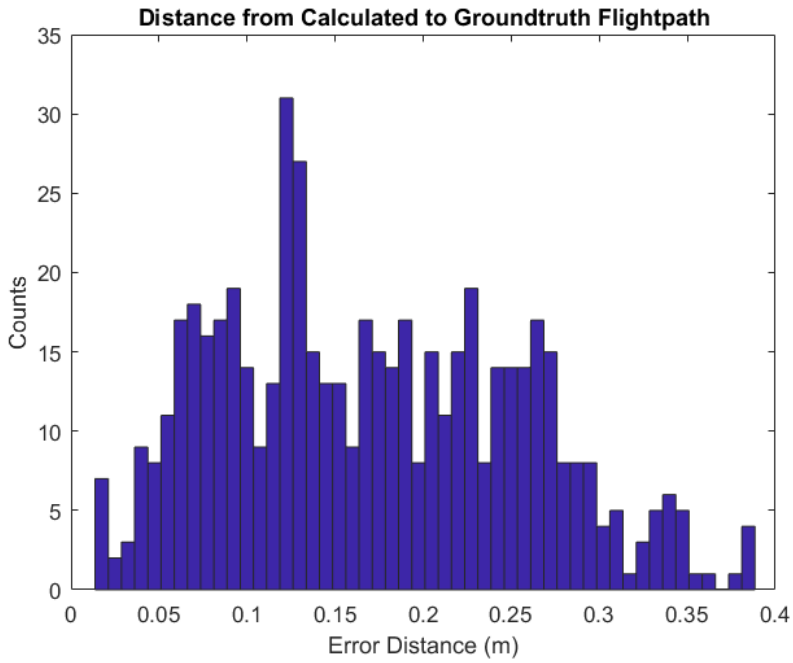


Figure 2.21: This figure illustrates the distance between the calculated to ground truth anchor positions.

8.3 2D Experiment

The main purpose of this experiment is to be able to visualise the convergence of the anchor positions and to be able to test the robustness of our algorithm. Due to the short range of the UWB devices and the obstacle of the Meeting Rooms, there was no instance of a given quadcopter position having full distance measurements to all 6 anchors. In this dataset there was 49.61% of the data missing. Here the (3,3) minimal solver was used.

Here the distance from the final calculated anchor position to the ground truth anchor positions are as follows, $(0.0814, 0.1717, 0.0702, 0.1392, 0.1700, 0.2017)m$, for anchors 1 to 6 respectively. Again a retrospective Procrustes fitting was applied as before.

9 Conclusions

In this paper, a novel method has been constructed to take advantage of the factorisation of the transmitter-receiver matrix in order to push for real-time Ultra-Wideband anchor calibration. This has been verified using TOA Ultra-Wideband measurement data in a streamed way to simulate real-time calculations.

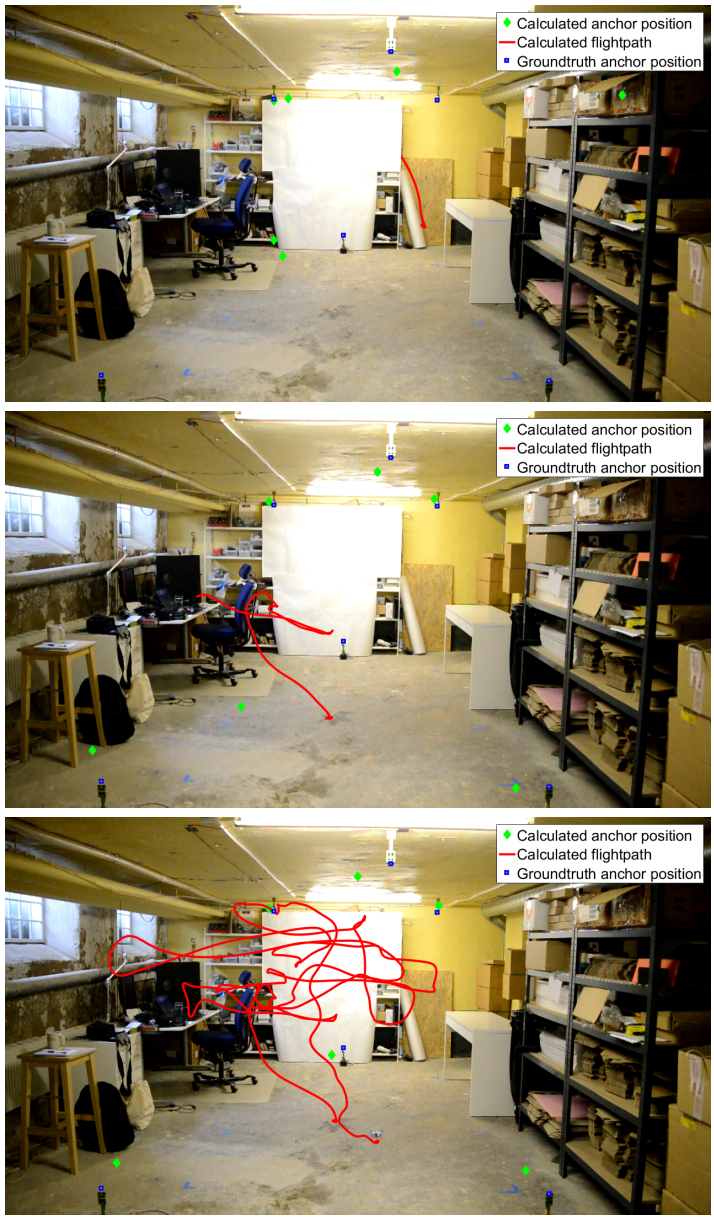


Figure 2.22: These figures illustrate the estimated anchor positions and the flightpath positions at various times of the flight to show how the anchor positions are updated. The positions of the anchors are shown by the blue square and the quadcopter can be seen in the photographs.

Looking at Figure 2.18, it can be seen that regardless of the buffer length, each solution converges to a similar solution. Once the system acquires a sufficient amount of data, they all converge. This then implies, by having a shorter buffer length then there would

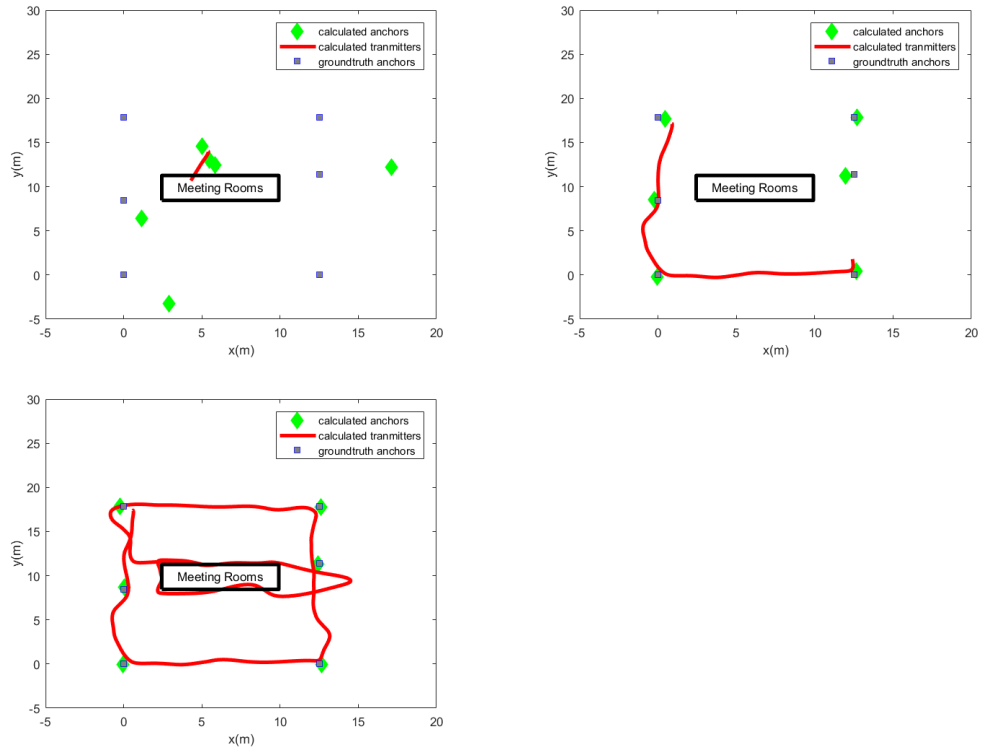


Figure 2.23: This figure illustrates the estimated anchor positions and the quadcopter positions. This is overlaid on the ground truth anchor positions. The central rectangle represents the Meeting rooms which provoked missing data. Each graph represents a different (increasing) time instance, the last being the final solution.

be more instances to test to see if a final solution has been met and therefore be able to move over to a computationally lighter positioning algorithm. It is also worth noting, that the there was a slight improvement in the final solution when the buffer length was smaller. Here the buffer length of 10 measurements gave the best solution. This may be due to higher frequency of optimisation steps, which would allow the calculated solutions to be closer to the global optimum, hence the solution to the next iteration would have a better initialisation, which is important in this highly non-linear system. Although a smaller buffer length tends to improve the final result, caution must be taken. The smaller the buffer length, the less robust our solver becomes so the errors from one section could affect the subsequent solutions and the final solution. This effect was seen on occasions, but UWB in line-of-sight situations produces accurate measurements so it didn't often occur and the buffer length was increased to 20 measurement to ensure robustness.

Looking at the results from the MOCAP studio experiments, Figures 2.19, 2.20 & 2.21, it can be seen that this method produces accurate results. For current Ultra-Wideband systems, the chip sets come with a recommended accuracy of $\pm 0.2m$. From our results we

are also able to achieve this accuracy, as seen from the convergent plots Figure 2.18 & 2.19. Further to this, each time we process the data, we reach a similar result. This would lead use to believe that we have achieved the global optima. It is also interesting to note that the convergence of the anchor positions to their true positions is not monotonic. It shows potentially at points in time that not enough data has been collected to make a reasonable solution but once it does have enough data, the convergence is rapid and the previous solutions are then corrected. This is further demonstrated in Figure 2.22.

In the 2D experiment, our algorithms were pushed in order to test the robustness of the system in the presence of large amounts of missing data. From Figure 2.23 it can be seen that the anchor positions are calculated to a high accuracy despite 49.61% of missing data. At most given quadcopter positions, it could acquire at most 4 distance measurements to the anchors. The minimum required number of overlapping anchors for the merging scheme to function is 2, any less and the 2 solution sets would be disjointed so it would be impossible to merge the coordinate systems. This data set, however, does show that our real-time scheme could be implemented on a large scale. As long as the criterion for the merging scheme is met then the system can be extend substantially. Only the local anchors to the quadcopter are needed for self-calibration of the anchors and relative positioning of the quadcopter. The previous solutions are only needed to have a common coordinate system. This then implies that if enough data is collected in a single area, then when the quadcopter moves to another area, only the anchor positions need to be stored and the data from that area would not be needed in further optimisation steps. This would allow for much faster computations and hybrid self-calibration/ trilateration algorithms.

For future work, the study of this convergent nature would be interesting in order to produce a viable method for a stopping condition for calibrating the anchor positions. By solving this problem it would allow for large scale navigation and positioning with no need for prior knowledge of the anchor positions, it would also be less computationally intensive which is preferable for robots and IoT devices.

Acknowledgements

The authors would like to thank the people at Bitcraze AB for all there support. The author would also like to that MAPCI, ELLIIT and Royal Physiographic Society of Lund for funding.

Bibliography

[1] B. Li, J. Salter, A. G. Dempster, and C. Rizos, "Indoor positioning techniques based on wireless lan," in *LAN, FIRST IEEE INTERNATIONAL CONFERENCE ON WIRELESS BROADBAND AND ULTRA WIDEBAND COMMUNICATIONS*, pp. 13–16.

[2] A. A. Adebomehin and S. D. Walker, "Enhanced ultrawideband methods for 5g los sufficient positioning and mitigation," in *2016 IEEE 17th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, pp. 1–4, June 2016.

[3] F. Dowla, F. Nekoogar, D. Benzel, G. Dallum, and A. Spiridon, "Method of remote powering and detecting multiple uwb passive tags in an rfid system," May 29 2012. US Patent 8,188,841.

[4] N. Priyantha, H. Balakrishnan, E. Demaine, and S. Teller, "Anchor-free distributed localization in sensor networks," in *Proceedings of the 1st international conference on Embedded networked sensor systems*, pp. 340–341, ACM, 2003.

[5] M. Oskarsson, K. Åström, and A. Torstensson, *Prime Rigid Graphs and Multidimensional Scaling with Missing Data*, pp. 750–755. IEEE–Institute of Electrical and Electronics Engineers Inc., 2014.

[6] G. Young and A. Householder, "Discussion of a set of points in terms of their mutual distances," *Psychometrika*, vol. 3, no. 1, pp. 19–22, 1938.

[7] P. Biswas, T.-C. Lian, T.-C. Wang, and Y. Ye, "Semidefinite programming based algorithms for sensor network localization," *ACM Trans. Sen. Netw.*, vol. 2, pp. 188–220, May 2006.

[8] T. Eren, O. Goldenberg, W. Whiteley, Y. Yang, A. Morse, B. Anderson, and P. Belhumeur, "Rigidity, computation, and randomization in network localization," in *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, vol. 4, pp. 2673–2684, IEEE, 2004.

[9] H. Stewénius, *Gröbner Basis Methods for Minimal Problems in Computer Vision*. PhD thesis, Lund University, APR 2005.

[10] Y. Kuang, S. Burgess, A. Torstensson, and K. Åström, "A complete characterization and solution to the microphone position self-calibration problem," in *The 38th International Conference on Acoustics, Speech, and Signal Processing*, 2013.

[11] S. Burgess, Y. Kuang, and K. Åström, "Node localization in unsynchronized time of arrival sensor networks," in *Proceedings of the 21st International Conference on Pattern Recognition*, 2012.

[12] S. Burgess, Y. Kuang, and K. Åström, "Pose estimation from minimal dual-receiver configurations," in *Proceedings of the 21st International Conference on Pattern Recognition*, 2012.

[13] M. Pollefeys and D. Nister, "Direct computation of sound and microphone locations from time-difference-of-arrival data," in *Proc. of International Conference on Acoustics, Speech and Signal Processing*, 2008.

[14] M. Crocco, A. Del Bue, M. Bustreo, and V. Murino, "A closed form solution to the microphone position self-calibration problem," in *37th International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2012), Kyoto, Japan*, March 2012.

[15] M. Crocco, A. Del Bue, and V. Murino, "A bilinear approach to the position self-calibration of multiple sensors," *Trans. Sig. Proc.*, vol. 60, pp. 660–673, feb 2012.

[16] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–95, 1981.

[17] C. Eckart and G. Young, "The approximation of one matrix by another of lower rank," *Psychometrika*, vol. 1, no. 3, pp. 211–218, 1936.

[18] T. Wiberg, "Computation of principal components when data are missing," in *Proc. Second Symp. Computational Statistics*, 1976.

[19] H. Aanaes, R. Fisker, K. Åström, and J. Carstensen, "Robust factorization," *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2002.

[20] A. M. Buchanan and A. W. Fitzgibbon, "Damped newton algorithms for matrix factorization with missing data," in *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 2 - Volume 02*, CVPR '05, (Washington, DC, USA), pp. 316–322, IEEE Computer Society, 2005.

[21] Q. Ke and T. Kanade, "Robust L_1 norm factorization in the presence of outliers and missing data by alternative convex programming," in *Proc. Conf. Computer Vision and Pattern Recognition*, 2005.

[22] A. Eriksson and A. Hengel, "Efficient computation of robust weighted low-rank matrix approximations using the L_1 norm," *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2012.

[23] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?," *Journal of ACM*, 2009.

[24] R. Garg, A. Roussos, and L. Agapito, "Dense variational reconstruction of non-rigid surfaces from monocular video," in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pp. 1272–1279, IEEE, 2013.

[25] C. Olsson and M. Oskarsson, "A convex approach to low rank matrix approximation with missing data," in *Scandinavian Conf. on Image Analysis*, 2011.

[26] R. Cabral, F. D. la Torre, J. P. Costeira, and A. Bernardino, "Unifying nuclear norm and bilinear factorization approaches for low-rank matrix decomposition," in *Proc. Int. Conf. on Computer Vision*, 2013.

[27] L. W. Mackey, M. I. Jordan, and A. Talwalkar, "Divide-and-conquer matrix factorization," in *Advances in Neural Information Processing Systems*, pp. 1134–1142, 2011.

[28] V. Larsson, C. Olsson, E. Bylow, and F. Kahl, "Rank minimization with structured data patterns," in *ECCV*, 2014.

[29] C. Julià, A. D. Sappa, F. Lumbreras, J. Serrat, and A. López, "An iterative multiresolution scheme for sfm with missing data," *Journal of Mathematical Imaging and Vision*, vol. 34, no. 3, pp. 240–258, 2009.

[30] V. Larsson and C. Olsson, "Convex envelopes for low rank approximation," in *EMMCVPR 2015*, 2015.

[31] F. Jiang, M. Oskarsson, and K. Åström, "On the minimal problems of low-rank matrix factorization," in *Proc. Conf. Computer Vision and Pattern Recognition*, 2015.

[32] K. Batstone, M. Oskarsson, and K. Åström, "Robust time-of-arrival self calibration and indoor localization using wi-fi round-trip time measurements," in *2016 IEEE International Conference on Communications Workshops (ICC)*, pp. 26–31, May 2016.

Paper VI

Collaborative Merging of Radio SLAM Maps in View of Crowd-sourced Data Acquisition and Big Data

Kenneth Batstone, Magnus Oskarsson and Kalle Åström
Centre of Mathematical Sciences, Lund University, Lund, Sweden

Abstract: Indoor localization and navigation is a much researched and difficult problem. The best solutions, usually use expensive specialized equipment and/or prior calibration of some form. To the average person with smart or Internet-Of-Things devices, these solutions are not feasible, particularly in large scales. With hardware advancements making Ultra-Wideband devices more accurate and low powered, this unlocks the potential of having such devices in commonplace around factories and homes, enabling an alternative method of navigation. Therefore, indoor anchor calibration becomes a key problem in order to implement these devices efficiently and effectively. In this paper, we present a method to fuse radio SLAM (also known as Time-Of- Arrival self-calibration) maps together in a linear way. In doing so we are then able to collaboratively calibrate the anchor positions in 3D to native precision of the devices. Furthermore, we introduce an automatic scheme to determine which of the maps are best to use to further improve the anchor calibration and its robustness but also show which maps could be discarded. Additionally, when a map is fused in a linear way, it is a very computationally cheap process and produces a reasonable map which is required to push for crowd-sourced data acquisition.

1 Introduction

Navigation has become a key part of modern civilisation, with most people using systems such as GPS on a daily basis, in their cars or on their person, integrated into their smart phones. The demand for positioning systems is also increasing with the era of 5G upon us. With 5G, we expect to see an increase of positioning accuracy in addition to having more devices, such as Internet-Of-Things (IoT), to also required positioning. For example, items in warehouses will require positioning to enable automation in the warehouses to improve efficiency.

Currently GPS provides good positioning for most users in an outdoor environment. Unfortunately, this cannot be said once inside a building. Once inside, the GPS signals are heavily attenuated, meaning the accuracy of the positioning can decrease to encompass a whole build or more. When this occurs users must use an alternative system to navigate indoors.

There are currently many options to overcome this problem but they all come with their own drawbacks. In robotics, many use optical devices to perform SLAM (Simultaneous localization and mapping, [1]) such as cameras and LIDAR, which produce good results but such devices can be expensive and computationally tasking. This restricts such methods to small environments with a low amount of dynamic features. For mobile phone users, a large focus has been using the signal strength of Wi-Fi networks to perform positioning since the infrastructure currently exists in most buildings but due to the nature of radio signals in complex environments, they have a low accuracy and with distance, the errors increase exponentially, [2].

One such technology which is commercially available is Ultra-Wideband (UWB). These devices are low powered and perform 2-way timing in order to obtain high precision in positioning, between two devices. This unlocks the potential of having such devices in common place around factories and homes, enabling an alternative method of navigation indoors for people and Internet of Things (IoT) devices.

Another technology which shows promise is round-trip time (RTT) being enabled on Wi-Fi. With RTT, it is expected to perform ranging between routers and mobile device with as low as 1 metre accuracy. Many modern routers have the ability to perform this but currently awaits a firmware update. This technology uses the 802.11mc IEEE standard, which has been enabled on Android Pie devices. A strong advantage to this option is that the infrastructure already exists.

With these developments comes further issues. Due to the large number of devices, calibration of the anchors becomes problematic. Currently large datasets require vast amounts of memory and processing power, which is impossible for most machines. In this paper we

present new research on methods for large scale anchor self-calibration problem. Here we present a method to merge maps together in a linear way. In doing so we build a library of tools in order to determine the quality of each map and then to be able to fuse them together to produce a global map. Additionally, when a map is merged in such a way, it is a relatively computationally cheap process and produces a reasonable map which is required to push for crowd-sourced data acquisition. The proposed method will help bridge the memory requirement issues and the ability to select the best datasets.

The proposed method was tested on both simulated and real UWB distance measurements. These datasets are created using 2-way timing, therefore it is a Time-Of-Arrival (TOA) self-calibration problem.

The TOA self-calibration problem, is the problem of determining the positions of a number of receivers and transmitters given only receiver-transmitter distances. Here, there is no prior knowledge of the anchor positions.

2 Background

To solve such problems, one method is to solve a minimal case and extend that solution. Minimal cases for low rank matrix factorization, for missing data, were investigated in [3]. In [4] a RANSAC paradigm was used in conjunction with minimal solvers and explored in order to obtain a robust and fast solution of the TOA self-calibration problem, with missing data and noise. In [5] a sequential merging scheme was created to explore the potential of real-time anchor calibration. One pitfall of the described scheme was that as more data was collected, memory requirements and computational complexity increased which limited the system.

The TOA self-calibration problem and other indoor SLAM methods are rarely performed in large scale using radio. Computer vision research has address some issues common to both optical SLAM and TOA self-calibration problem, such as memory limits, accuracy and computational power. In [6] and [7], the authors exploit the structure of the Jacobian so that memory limits and computational complexity are improved to allow for SLAM in larger environments with acceptable losses in accuracy.

In [8], a solution was given for large scale SLAM, with promising memory requirements, computational effort and an accuracy of $0.5m$ in 2D, but this works differs since the authors use foot mounted inertial measurement units (IMU) to crowd-source SLAM maps, which is not as prevalent as radio infrastructure. In [9] map fusion was explored for a multi-robot SLAM framework, but this method was tested on only two maps. More research has been conducted in this area, [10, 11], but still very few robots and maps are used when merging. In [12], the authors address the issue of large memory requirements needed for Collaborative

Visual SLAM. Although optical SLAM and TOA self-calibration share similar solutions to similar problems, they differ greatly in accuracy and the type of data. In optical SLAM many other instruments on the robot assist the formation of the solution and improves the accuracy. This provides a rich and reliable dataset. For the TOA self-calibration using radio systems, it is common that there are fewer anchor positions than user sender positions. This means that when merging anchor positions, the sparsity of the data is a constraint on the solution and prone to errors due to the accuracy of the ranging.

3 Method

We will now describe the basic underlying geometry of our problem. Let $\mathbf{R}_i$, $i = 1, \ldots, m$ and $\mathbf{S}_j$, $j = 1, \ldots, n$ be the spatial coordinates of m receivers (e.g. Ultra-Wideband anchors) and n transmitters (e.g. Crazyflie quadcopter), respectively. For measured time of arrival t_{ij} from transmitter $\mathbf{R}_i$ and receiver $\mathbf{S}_j$, we have $vt_{ij} = \|\mathbf{R}_i - \mathbf{S}_j\|_2$, where v is the speed of measured signals and $\|.\|_2$ is the ℓ^2-norm. The speed v is assumed to be known and constant. We further assume that we, at each receiver can distinguish which transmitter j each event is originating from. This can be done e.g. if the signals are temporally separated or using different frequencies. We will in the following work with the distance measurements $d_{ij} = vt_{ij}$. It is quite common that such data contains both missing data from poor signal communications and outliers due to inaccuracies of the hardware measurements. The TOA calibration problem can then be defined as follows,

Problem 5. *(Time-of-Arrival Self-Calibration) Given absolute distance measurements*

$$d_{ij} = \|\mathbf{R}_i - \mathbf{S}_j\|_2 + \epsilon_{i,j}, \tag{2.57}$$

where the receiver positions are defined as $\mathbf{R}_i$, $i = 1, \ldots, m$ and transmitter positions as $\mathbf{S}_j$, $j = 1, \ldots, n$. Here the errors $\epsilon_{i,j}$ are assumed to be either inliers, in which case the errors are small ($\epsilon_{i,j} \in N(0, \sigma)$) or outliers, in which case the measurements are way off.

Here we will use the set W for the indices (i, j) corresponding to the inlier measurements.

The Time-of-Arrival Self-Calibration problem can be solved by computing the bundle adjustment of (2.58), shown below.

$$\min_{\mathbf{R},\mathbf{S}} \sum_{(i,j)\in W} (d_{i,j} - ||\mathbf{R}_i - \mathbf{S}_j||_2)^2. \tag{2.58}$$

For simplification, (2.58) can be represented as,

$$\text{argmin} \sum |d - f(\mathbf{R}, \mathbf{S})|^2 \tag{2.59}$$

where $f(\mathbf{R},\mathbf{S}) = ||\mathbf{R}-\mathbf{S}||_2$ can be the nonlinear function for all combinations of $\mathbf{R},\mathbf{S}$. Therefore it can be assumed that there exists an optimal $\mathbf{R}^\star$ and $\mathbf{S}^\star$, such that a_{opt} is the minima, ie.

$$a_{opt} = \sum |d - f(\mathbf{R}^\star,\mathbf{S}^\star)|^2. \tag{2.60}$$

Then the sum of the residuals can be linearized around $\mathbf{R}^\star$ and $\mathbf{S}^\star$ as

$$a(\mathbf{R},\mathbf{S}) \approx a_{opt} + J\begin{bmatrix}\mathbf{R}-\mathbf{R}^\star\\ \mathbf{S}-\mathbf{S}^\star\end{bmatrix}, \tag{2.61}$$

where J is the jacobian of f wrt. $\mathbf{R}$ and $\mathbf{S}$. The problem can be reformulated as

$$a(\mathbf{R},\mathbf{S}) \approx a_{opt} + \begin{bmatrix}J_\mathbf{R} & J_\mathbf{S}\end{bmatrix}\begin{bmatrix}\Delta\mathbf{R}\\ \Delta\mathbf{S}\end{bmatrix}. \tag{2.62}$$

We can include the contribution of $\mathbf{S}$ into the expression by solving

$$\min_{\Delta\mathbf{S}} \quad |a_{opt} + J_\mathbf{R}\Delta\mathbf{R} + J_\mathbf{S}\Delta\mathbf{S}|^2 \tag{2.63}$$

which has the closed form solution $\Delta\mathbf{S} = -(J_\mathbf{S}^T J_\mathbf{S})^{-1}J_\mathbf{S}^T J_\mathbf{R}\Delta\mathbf{R}$ where ^{T} denotes matrix transpose. Insertion into (2.62) yields

$$a(\mathbf{R},\mathbf{S}) \approx a_{opt} + (I - J_\mathbf{S}(J_\mathbf{S}^T J_\mathbf{S})^{-1}J_\mathbf{S}^T)J_\mathbf{R}\Delta\mathbf{R}, \tag{2.64}$$

where I denotes the identity matrix of proper size. In order to reduce the amount of data being saved in a database, a matrix A is introduced such that

$$A = (I - J_\mathbf{S}(J_\mathbf{S}^T J_\mathbf{S})^{-1}J_\mathbf{S}^T)J_\mathbf{R}. \tag{2.65}$$

Now A above can be decomposed into $A = VU$ where U is a upper triangular matrix and V is a unitary matrix. Hence,

$$\begin{aligned}|a(\mathbf{R},\mathbf{S})|^2 =& |V^T a(\mathbf{R},\mathbf{S})|^2 \approx |V^T a_{opt} + U\Delta\mathbf{R}|^2\\ =& |a_{opt}|^2 + |U\Delta\mathbf{R}|^2.\end{aligned} \tag{2.66}$$

If two such solutions are available, then the sum of the norms can be formulated as

$$\begin{aligned}&|a_1(\mathbf{R},\mathbf{S})|^2 + |a_2(\mathbf{R},\mathbf{S})|^2 \approx |a_{1,opt}|^2 + ...\\ &... + |U_1(\mathbf{R}-\mathbf{R}^\star{}_1)|^2 + |a_{2,opt}|^2 + |U_2(\mathbf{R}-\mathbf{R}^\star{}_2)|^2.\end{aligned} \tag{2.67}$$

This expression can be minimized for R as

$$\mathbf{R_{opt}} = (U_1^T U_1 + U_2^T U_2)^{-1}(U_1^T U_1\mathbf{R}^\star{}_1 + U_2^T U_2\mathbf{R}^\star{}_2). \tag{2.68}$$

Which has the general expression, for k maps as,

$$\begin{aligned}\mathbf{R_{opt}} =&(U_1^T U_1 + U_2^T U_2 + ... + U_k^T U_k)^{-1} \\ &(U_1^T U_1 \mathbf{R^\star}_1 + U_2^T U_2 \mathbf{R^\star}_2 + ... + U_k^T U_k \mathbf{R^\star}_k).\end{aligned} \tag{2.69}$$

Since in reality, some of the calculated maps will be erroneous due to the environment in which the measurements are taken, a weighting term is therefore introduced, where $\lambda_k \in [0, 1]$.

$$\begin{aligned}\mathbf{R_{opt}}(\lambda) = &(\lambda_1^2 U_1^T U_1 + \lambda_2^2 U_2^T U_2 + ... + \lambda_k^2 U_k^T U_k)^{-1} \\ &(\lambda_1^2 U_1^T U_1 \mathbf{R^\star}_1 + \lambda_2^2 U_2^T U_2 \mathbf{R^\star}_2 + ... + \lambda_k^2 U_k^T U_k \mathbf{R^\star}_k).\end{aligned} \tag{2.70}$$

In order to solve for the problem in (2.70), a new variant of the objective function for the full bundle adjustment (2.58) is used. Here, one only needs to solve for the vector $\lambda = [\lambda_1, ..., \lambda_k]$ as shown in (2.71),

$$\min_{\lambda} \sum_{k} \sum_{(i_k, j_k) \in \tilde{W}_k} (\lambda_k (d_{i_k j_k} - ||\mathbf{R_{opt}}_{i_k}(\lambda) - \mathbf{S}_{j_k}||_2))^2. \tag{2.71}$$

This, therefore, can be seen as a relaxation of (2.58), where the λ variable is similar to the weights in a weighted optimization. Due to the non-linearity of the problem, a good initialization is also needed for (2.71). To achieve this, a RANSAC scheme was devised to provide a good initialization but also an indication of which dataset are best to use, see Algorithm 6.

Algorithm 6 Our RANSAC Merging Scheme

1: Select 5 random maps
2: Calculate the optimal anchor positions using our merging algorithm.
3: **If:** The score of the objective function is the lowest value so far, select all maps within a $1m$ of RMSE distance of the optimal anchor positions. The initial 5 maps keep their λ values from the optimization, all other inliers are given a value of 0.5 and outliers are given a value of 0.
4: Repeat steps 1-3 200 times
5: Recalculate a new optimal anchor position using our merging algorithm with best lambda values as an initial guess to the optimization.

In this scheme, some of the values are arbitrary and can be tuned depending on the data type etc. These values are the 5 random maps and the selection of all maps within $1m$ of RMSE distance of the optimal anchor positions. The reason 5 maps were chosen is to maintain robustness, since the quality of maps vary, by choosing 5 maps the optimization can quickly determine a valid optimal anchor position for the majority of the iterations. The selection radius was chosen as generous catchment zone for the inlier set, this can be tuned to the specific need of the datasets.

4 Experimental Setup

4.1 Simulated Datasets

In order to test our method, three experiments were devised. The first experiment was to create 40 anchor positions and 1000 sender positions, randomly to span a $20 \times 20 \times 20m$ space. From there the distance matrix was calculated and Gaussian noise was added with a variance of $0.18m$ to simulate UWB measurements. A full bundle adjustment was then performed, in order to give a comparison to current state of the art method, [4]. The distance matrix, d, was then divided into 50 equal parts of 40 anchor positions and 20 sender positions and a map was created for each set. Then for each of the 50 maps, our method was tested with different optimization methods. Firstly our linear method (2.69), secondly our linear method with a weighting factor (2.70) and lastly a bundle adjustment. This was then repeated 400 times.

The second experiment was to perform the same experiment as above but to falsify 30% and 60% of the 50 maps. For the specific percentage of the maps, the anchor positions were randomly perturbed in a $40cm$ radius. The anchor positions were then transformed to ensure that the first anchor is the origin of the coordinate system and the second anchor on the x-axis and so on. This transformation was also performed on the sender positions. This therefore, would give a realistic poor result for those maps. The RANSAC method was then tested on these datasets to give an understanding of the robustness of the proposed merging schemes and to show how it could be used to determine good maps. Again this was repeated 400 times.

The third experiment was to test how the number of maps affects the time it takes to calculate the optimal solution. Once again, 40 anchor positions and 1000 sender positions were used like in the first experiment to simulate UWB measurements. The RANSAC method was then used to find a solution for different number of maps, with the time it took noted, and the time for a full bundle adjustment was also noted. This was only iterated once.

4.2 Real Datasets

For the final experiment, our algorithms were tested on real UWB measurements from a Bitcraze, Crazyflie quadcopter mounted with a UWB device (Decawave DWM1000 chip) in order to determine if the proposed method is feasible in a real world situation, shown in Figure 2.24. The experiment was conducted in a Motion Capture (MOCAP) Studio to give ground truth positions to compare our results with. There were 9 separate datasets with 6 anchors in the same position for each. The ground truth anchor positions were calculated

using the MOCAP cameras to a precision of ± 1 *mm*.

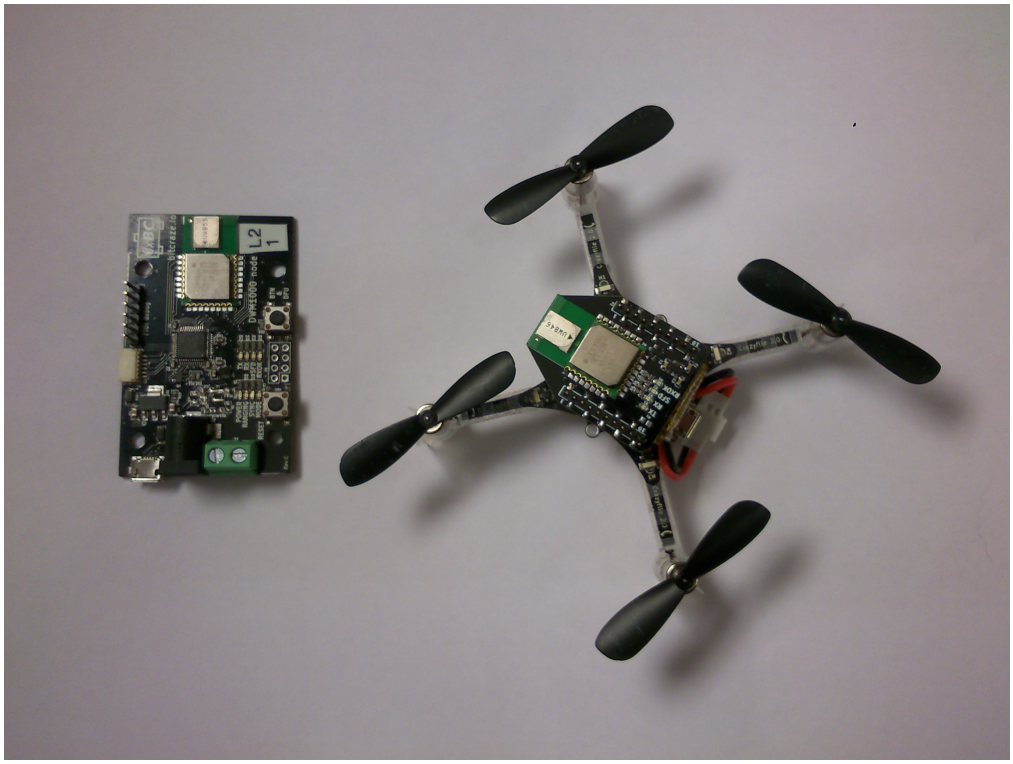


Figure 2.24: Image of the Ultra-Wideband anchor and Bitcraze Crazyflie quadcopter respectively from left to right.

Distance measurements from the quadcopter to all the anchors were measured at a frequency of 30 *Hz*.

The experiments were conducted by moving the quadcopter, by hand, around the room. The distance measurements were recorded so that they may be processed offline. Our algorithms do not require any prior knowledge of anchor or quadcopter positions. The only requirement is that the minimal solver (5,5) is satisfied for the 3D cases.

5 Results and Analysis

5.1 Simulated Datasets

In Figures 2.25 and 2.26 the results for the first experiment are shown. For all the experiments the Root Mean-Squared Errors (RMSE) are a comparison of the calculated optimal anchor positions to the ground truth anchor positions. It can been seen that each of the

methods perform differently, with the linear merging scheme being the least successful. The other three methods presented show good results since all three have at least 90% of the solutions with an error under the UWB distance measurement accuracy of $0.18m$, see Figure 2.26. The full bundle adjustment result was expected to be a good solution since the optimization is minimizing the residual for all 40000 distance measurements. Interestingly, the result for the RANSAC scheme did not achieve as good as a result for the best full bundle adjustment solutions but for 55% of the solutions it did achieve a better result. Furthermore, it has a steep curve at $0.03m$ RMSE distance Error. This indicates that the method reliably produces a similar result.

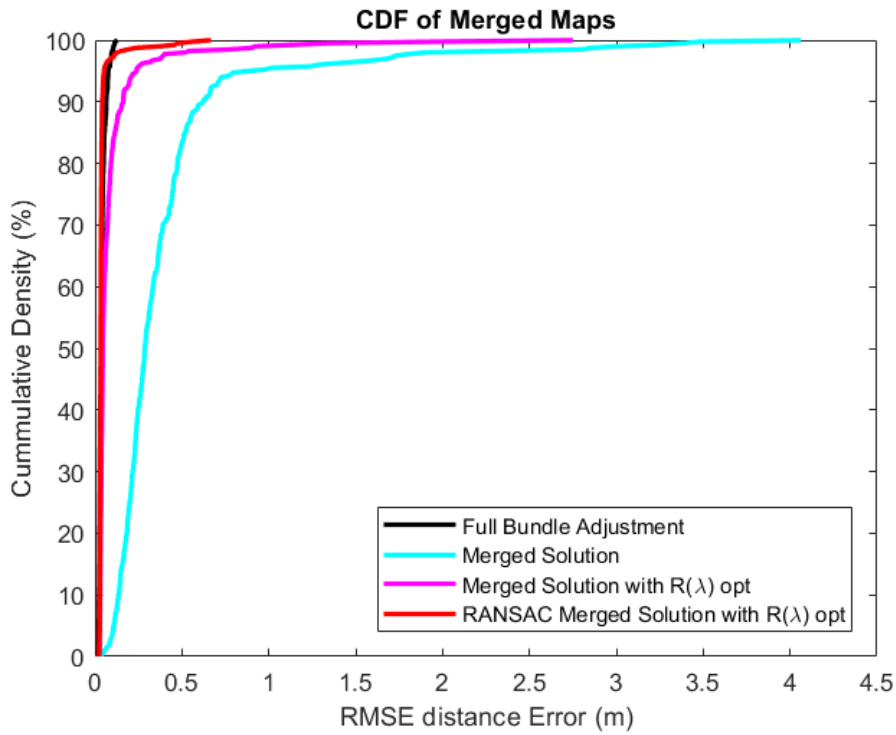


Figure 2.25: This figure illustrates the RMSE error for the each of the merged maps plotted against its cumulative density for 400 experiments.

In Figures 2.27 and 2.28, the results for the second experiment are shown. All three solutions show a similar result, which indicates that by using the RANSAC scheme it maintains its robustness. This is due to the RANSAC scheme being able to select a collection of maps which have similar and good results. This behaviour is further shown in Figure 2.28. Figure 2.28 is an example of the lambda values obtained after using the RANSAC scheme and the merging scheme with lambda optimization. It can be seen that the RANSAC scheme focuses the optimization of lambda in one cluster of the maps. This in turn produces better

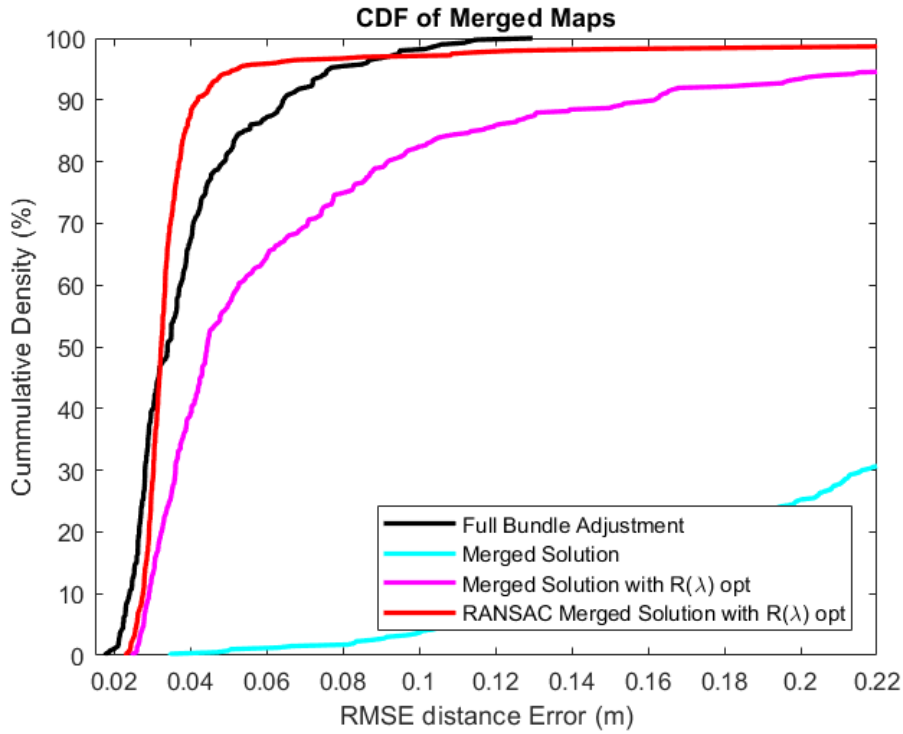


Figure 2.26: This figure is the same as Figure 2.25. It illustrates the RMSE error for the each of the merged maps plotted against its cumulative density for 400 experiments but only shows the RMSE range of 0.01 to 0.22.

optimal anchor positions, the RANSAC scheme had a RMSE distance error of $0.0379m$ and the merging scheme with lambda optimization $0.0498m$.

In Figure 2.29, the results for the third experiment is shown. It can be see that the time it takes to find a solution is dependent on the number of maps. Thus it also shows that the RANSAC method proposed is computationally cheaper. In this case, the trend appears to be parabolic which implies that there is an optimal number of maps for each experiment. This was computed for one random experiment, the times do vary for each experiment but the trends are similar.

5.2 Real Dataset

In Figure 2.30 are the lambda values obtained after using the RANSAC scheme and the merging scheme with lambda optimization. For this experiment the RMSE is a comparison of the calculated optimal anchor positions to the ground truth anchor positions. The RANSAC scheme had a RMSE distance error of $0.106m$ and the merging scheme with

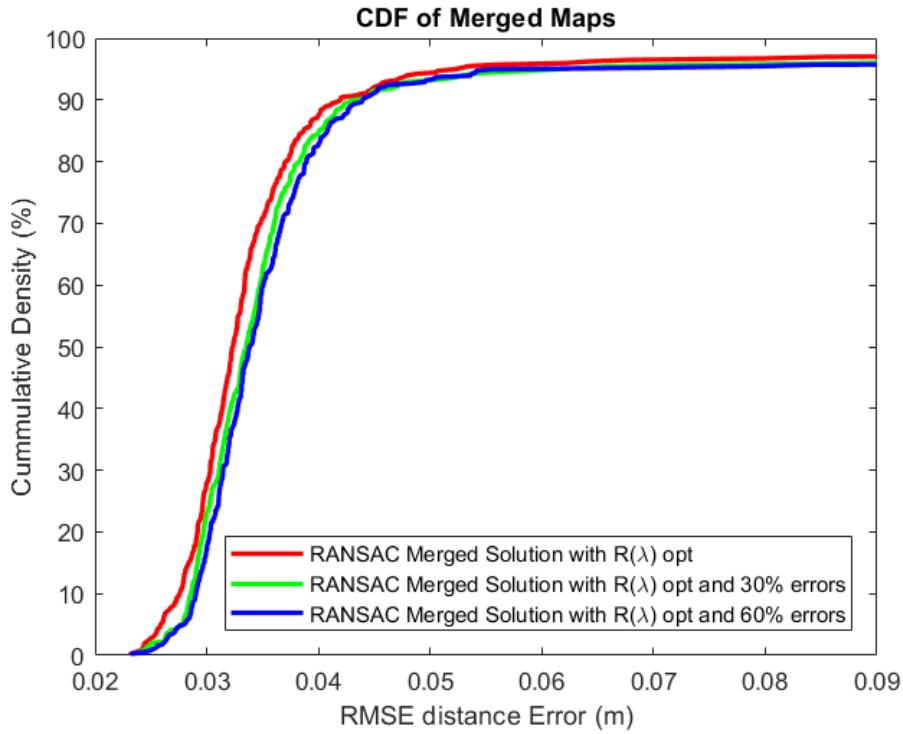


Figure 2.27: This figure illustrates the RMSE error for the each of the merged maps with different percentages of errors plotted against its cumulative density for 400 experiments.

lambda optimization $0.1369m$. Due to the restricted number of maps in this case, it is difficult to determine which of the schemes is better, since nearly all maps are needed to calculate an optimal map then the lambda value are similar.

6 Conclusions

In this paper, a method has been constructed to merge maps together in a linear way. In doing so we build a library of tools to determine the quality of each map, and once the quality of multiple maps were determined, we can logically merge them together to produce a global map.

Looking at the results from the MOCAP studio experiments, in Figure 2.30 it can be seen, that this method produces accurate results. For current Ultra-Wideband systems, the chip sets come with a recommended accuracy of ± 0.2m. From our results, we are also able to achieve this accuracy. It is also interesting to note that the lambda values for each of the

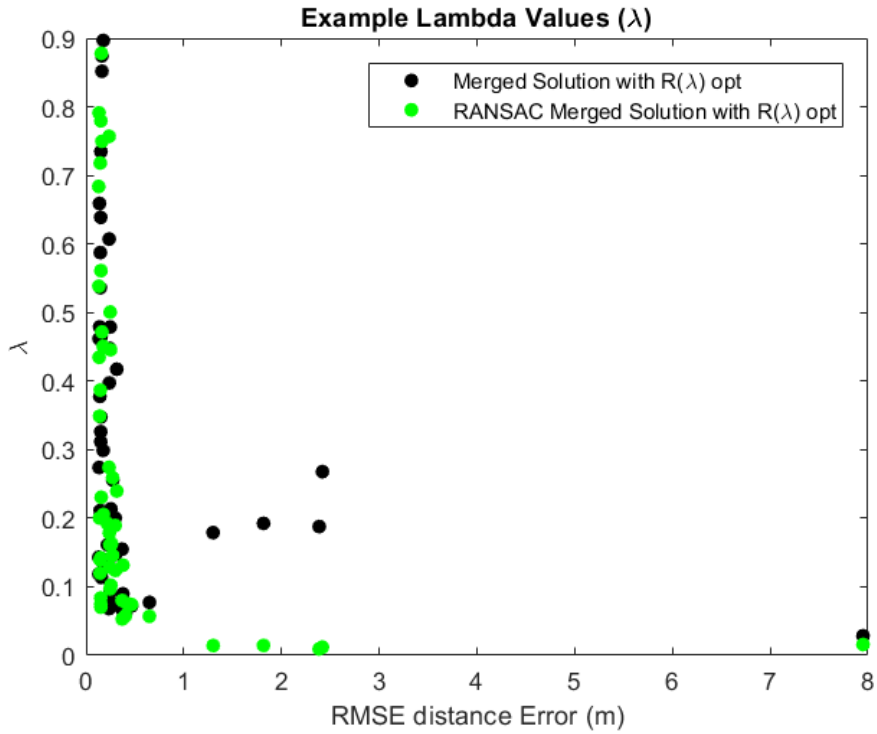


Figure 2.28: This figure illustrates the RMSE error for each map plotted against its calculated lambda value.

maps are varied, in particular the map with the smallest error has a lambda value of ca. 0.35. It shows potentially that there are not enough separate maps that have been collected to make a reasonable estimate of the quality of each map. One would expect the lambda values to decrease as the RMSE error increases, as seen in Figure 2.28, but on this occasions there are erroneous lambda values. This may be due to the non-linearity of the self-calibration problem, since there will be many local minimas, contributions from all maps may be used.

In the first experiment, our algorithms were pushed, to test the robustness of the system. From Figure 2.25, it can be seen that the anchor positions are calculated to a high accuracy in comparison to the full bundle adjustment. It can also be noted that roughly 98% of the merged maps had a RMSE error under $0.25m$. Of course, the full bundle adjustment produces a better result and is considered the gold standard but in reality it is not a viable option. The bundle adjustment is very computationally expensive and is limited by the size of the distance matrix. During the optimization of the bundle adjustment, it has to estimate 120000 parameters (40 anchors, 1000 senders, 3 degrees of freedom), which modern computers with large RAM can calculate but any larger wouldn't be possible. By partitioning the distance matrix, multiple solutions can be created in parallel, then merged

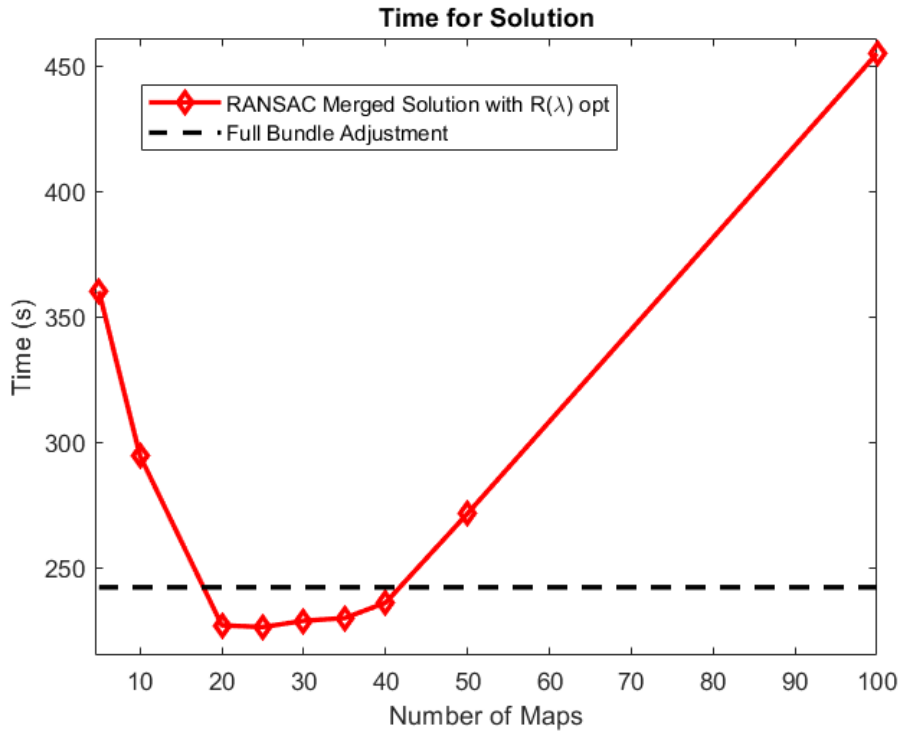


Figure 2.29: This figure illustrates the computatuional time for an optimal solution to be found for different number of maps. The time for a full Bundle Adjustment over all anchors and senders is also shown as a comparison.

together. In addition to this, once the lambda values have been calculated, one would have an estimate of the quality of each map and the ability to logically manage each solution with data storage and merged maps quality in mind. Another benefit, is that the number of parameters is reduced considerably when performing the merging algorithm with weights. In this case from 120000 parameters to 50 parameters.

The main advantage with such linear fusion is that it is a relatively computationally cheap process, that unlocks the potential for crowd-sourced data acquisition without compromising map quality. In our case, for the simulated dataset with 60 % errors, to perform a bundle adjustment on all 50 maps and merge them took 47 minutes, whereas the full bundle adjustment took 1.5 days on the same machine for all 400 iterations. This can be seen further in Figure 2.29, with an appropriate number of maps. Although the computational time reduction can be seen, it is not as large as the one mentioned for the simulated dataset with 60 % errors. This may be due to the the RANSAC initialization, this step produces a robust and close initialization which reduces the time needed of our method to converge to the optimal solution. In the case for the simulated data used in experiment 3, since all the maps are viable (no outliers) then many more maps are initialized with the value 1, hence

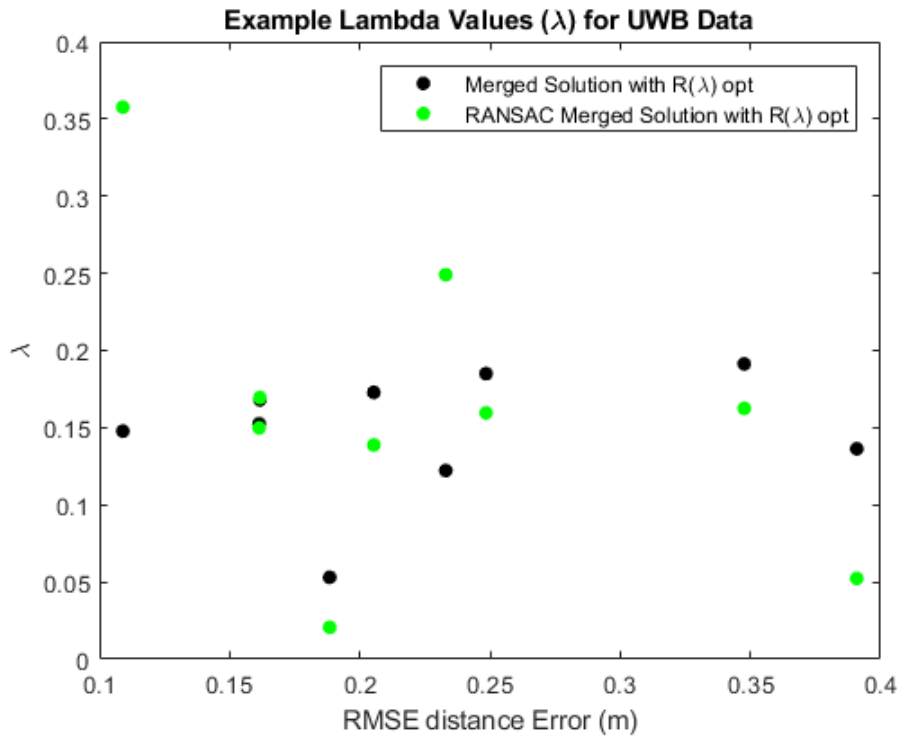


Figure 2.30: This figure illustrates the RMSE error for map plotted against its calculated lambda value. The maps are created using UWB mounted on a quadcopter.

the computational time is less affected. The proposed method bridges memory requirement issues and offers the ability to select the best datasets. In addition to this, the method would also work for different media type, such as bluetooth, multiple WiFi frequencies and optical SLAM. Provided that the positions of the anchor points are the same for each media.

For future work, the study of a collaborative data management scheme would be highly advantageous. In doing so, would give an autonomous way of choosing which parts of the dataset to fuse in order to discard unnecessary data and keep only the required data to improve a map. For instance, if an office building were to be mapped using crowd-sourced data, there would exist areas that would be oversampled, such as the main entrance and corridors. Whereas a storage room would be sampled infrequently, therefore an automatic scheme that would discard the oversampled areas would be advantageous to data management. In summary, this would be a way of determining the uniqueness of a given map.

Bibliography

[1] H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping: part i," *IEEE Robotics Automation Magazine*, vol. 13, pp. 99–110, June 2006.

[2] B. Li, J. Salter, A. G. Dempster, and C. Rizos, "Indoor positioning techniques based on wireless lan," in *LAN, FIRST IEEE INTERNATIONAL CONFERENCE ON WIRELESS BROADBAND AND ULTRA WIDEBAND COMMUNICATIONS*, pp. 13–16.

[3] F. Jiang, M. Oskarsson, and K. Åström, "On the minimal problems of low-rank matrix factorization," in *Proc. Conf. Computer Vision and Pattern Recognition*, 2015.

[4] K. Batstone, M. Oskarsson, and K. Åström, "Robust time-of-arrival self calibration and indoor localization using wi-fi round-trip time measurements," in *2016 IEEE International Conference on Communications Workshops (ICC)*, pp. 26–31, May 2016.

[5] K. Batstone, M. Oskarsson, and K. Åström, "Towards real-time time-of-arrival self-calibration using ultra-wideband anchors," in *2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pp. 1–8, Sept 2017.

[6] M. Byröd and K. Åström, *Bundle Adjustment using Conjugate Gradients with Multiscale Preconditioning*. 2009.

[7] M. Byröd and K. Åström, *Conjugate Gradient Bundle Adjustment*, vol. 6312, pp. 114–127. Springer, 2010.

[8] M. G. Puyol, P. Robertson, and M. Angermann, "Managing large-scale mapping and localization for pedestrians using inertial sensors," in *2013 IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*, pp. 121–126, March 2013.

[9] F. Chanier, P. Checchin, C. Blanc, and L. Trassoudaine, "Map fusion based on a multi-map slam framework," in *2008 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, pp. 533–538, Aug 2008.

[10] P. Schmuck and M. Chli, "Multi-uav collaborative monocular slam," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3863–3870, IEEE, 2017.

[11] S. Liu, K. Mohta, S. Shen, and V. Kumar, "Towards collaborative mapping and exploration using multiple micro aerial robots," in *Experimental Robotics*, pp. 865–878, Springer, 2016.

[12] D. Van Opdenbosch, T. Aykut, N. Alt, and E. Steinbach, "Efficient map compression for collaborative visual slam," in *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 992–1000, IEEE, 2018.

Paper VII

Robust Phase-Based Positioning Using Massive MIMO With Limited Bandwidth

Xuhong Li[1], Kenneth Batstone[2], Kalle Åström[2], Magnus Oskarsson[2], Carl Gustafson[1], Fredrik Tufvesson[1]
[1] *Dept. of Electrical and Information Technology, Lund University, Lund Sweden*
[2] *Centre of Mathematical Sciences, Lund University, Lund Sweden*

Abstract: This paper presents a robust phase-based positioning framework using a massive MIMO system. The phase-based distance estimates of MPCs together with other parameters are tracked with an EKF, the state dimension of which varies with the birth-death processes of paths. The RIMAX and the modeling of dense multipath component in the framework further enhance the quality of parameter tracking by providing an accurate initial state and the underlying noise covariance. The tracked MPCs are fed into a time-of-arrival self-calibration positioning algorithm for simultaneous trajectory and environment estimation. Throughout the positioning process, no prior knowledge of the surrounding environment and base station position is needed. The performance is evaluated with the measurement of a 2D complex movement, which was performed in a sports hall with an antenna array with 128 ports as base station using a standard cellular bandwidth of 40 MHz. The positioning result shows that the mean deviation of the estimated user equipment trajectory from the ground truth is 13 cm. In summary, the proposed framework is a promising high-resolution radio-based positioning solution for current and next generation cellular systems.

1 Introduction

High precision positioning information is a fundamental component of autonomous systems and location-aware applications in mobile devices. To pursue better user experience, these new services and systems keep bringing new challenges to the positioning systems regarding the accuracy, reliability, etc. The Global Positioning System (GPS) works well outdoors, but the accuracy and robustness degrade severely in scenarios like urban canyons and indoor environments due to poor propagation conditions between satellites and user equipment (UE). In contrast, cellular and wireless networks generally have good coverage in those GPS harsh environments. As a substitute or supplement to GPS, much effort has been put into the research of radio-based positioning techniques.

Accurate radio-based positioning commonly relies on geometrical information (distance, delay and angle) of multipath components (MPCs) from the radio channel. The estimation quality of these channel parameters in turn determines the positioning performance. In recent years, positioning with ultra-wideband (UWB) signal has drawn special interest because of the excellent accuracy[1]. The fine delay resolution due to the large bandwidth used makes it possible to resolve MPCs and track the distance changes in centimeter level. However, UWB positioning can only be applied in limited scenarios considering it is a low-power and short-range technique. These shortcomings naturally lead us to the question: is it possible to deliver comparable positioning accuracy by utilizing limited bandwidth in both indoor and outdoor scenarios? We try to solve the puzzle from a channel modelling perspective. The wireless propagation channel is commonly characterized as a sum of specular-like paths and non-resolvable components. Considering that cellular systems are typically operating with a carrier frequency at a few GHz with a bandwidth of 20-40 MHz, the delay resolution is in a scale of 7.5-15 m. However, we notice that the corresponding wavelengths are in the order of centimeters and one wavelength corresponds to a 2π phase shift. For each MPC, the delay and phase are two parameters which vary simultaneously with the wave propagating. If the spatial sampling rate of the radio channel is sufficiently high, i.e., taking a few snapshots within one wavelength movement, it is possible to track the distance change in centimeter level by measuring the phase shift between two consecutive snapshots. With limited bandwidth, the coherence in the delay domain is a challenge for successfully detecting and tracking many MPCs simultaneously. However, the large-scale antenna array could provide additional distinction between MPCs in the spatial domain. The feasibility of the phase-based positioning has been preliminarily proved in [2]. In that work, the phase and delay are assumed to be independent parameters, which are estimated separately and only the phase is used for the movement tracking. Because phase and delay affect each other with wave propagating, and phase estimates are usually discontinuous in complex environments, there are risks of causing errors or losing tracks.

Motivated by the above analysis, we present a robust phase-based positioning framework

in this paper. As a proof-of-concept study, the focus of this work is on demonstrating the possibility of high-resolution radio-based positioning given limited bandwidth and with no prior environment knowledge, rather than on reducing system complexity to implement real-time positioning. Based on some well-established algorithms, e.g,. the Extended Kalman Filter (EKF) [3] and the iterative maximum-likelihood estimation algorithm (RIMAX) [4], the MPC parameters are extracted from the channel measurement data. A time-of-arrival (TOA) self-calibration positioning algorithm, which is a structure-of-motion approach and widely used in image processing, is finally applied for simultaneous UE trajectory and environment estimation. The main contributions are

- The unique mapping between the phase shift and the distance change in our dynamic model leads to a simpler kinematic model by involving less parameters, and the robustness of the system is also improved.
- The performance is evaluated with real measurements of a complex movement. The results prove that the proposed framework provides outstanding MPC tracking and positioning performance even with limited bandwidth.

The paper is structured as follows. In Section II, dynamic propagation channel modeling is discussed. Section III introduces the estimation of path parameters with the EKF. Section IV describes details of the measurement campaign. The MPC tracking results are presented in Section V. Section VI introduces the TOA positioning algorithm and positioning result is presented. Finally, Section VII concludes the paper.

2 Dynamic Propagation Channel Modeling

An observation of the propagation channel, the impulse response $\mathbf{h}_k$ could be decomposed into three non-overlapping components: specular components $\mathbf{h}_{sp}$, dense multipath component (DMC) $\mathbf{h}_{dmc}$ and measurement noise $\mathbf{h}_n$, yielding

$$\mathbf{h}_k = \mathbf{h}_{sp} + \mathbf{h}_{dmc} + \mathbf{h}_n. \tag{2.72}$$

Positioning relies on the geometrical information from $\mathbf{h}_{sp}$, which is characterized as a superposition of MPCs. The other two components constitute measurement impairments for our purpose.

2.1 Channel Model

In the proposed framework, the double-directional radio channel model [5] is employed to extract the spatial and temporal information of the MPCs from the measured channel

transfer function $\mathbf{H} \in \mathbb{C}^{N_s \times N_f \times N_{\mathrm{Tx}} \times N_{\mathrm{Rx}}}$, given as

$$\mathbf{H}(f) = \sum_{l=1}^{L} \boldsymbol{\gamma}_l e^{-j2\pi f \frac{d_l}{c}} \mathbf{G}_{\mathrm{Rx}}(\varphi_{\mathrm{Rx},l}, \theta_{\mathrm{Rx},l}) \mathbf{G}_{\mathrm{Tx}}(\varphi_{\mathrm{Tx},l}, \theta_{\mathrm{Tx},l})^{\mathrm{T}}, \tag{2.73}$$

where N_s, N_f, N_{Tx} and N_{Rx} refer to the number of channel snapshots, frequency sample points, transmit and receive antenna elements. $\mathbf{G}_{\mathrm{Tx}} \in \mathbb{C}^{N_{\mathrm{Tx}} \times N_a N_e}$ and $\mathbf{G}_{\mathrm{Rx}} \in \mathbb{C}^{N_{\mathrm{Rx}} \times N_a N_e}$ describe the far-field antenna response of all antenna array ports at the transmit and receive sides, with respect to the azimuth and elevation angles of departure (AODs) $(\varphi_{\mathrm{Tx},l}, \theta_{\mathrm{Tx},l})$ and angles of arrival (AOAs) $(\varphi_{\mathrm{Rx},l}, \theta_{\mathrm{Rx},l})$ of the lth path. N_a and N_e represent the number of azimuth and elevation angular samples. L is the number of propagation paths. The complex path weight is parametrized as $\boldsymbol{\gamma}_l = \boldsymbol{\alpha}_l e^{j\boldsymbol{\phi}_l}$ where $\boldsymbol{\alpha}_l$ and $\boldsymbol{\phi}_l$ represent the vectors of magnitude and phase, respectively. Instead of using delay, we directly interpret the phase shift as a distance measure, i.e., phase-based distance d_l. The time-variant structural vectors associated with the propagation environment geometry and the path weights are defined as

$$\mu = [d^{\mathrm{T}} \quad \boldsymbol{\varphi}_{\mathrm{Tx}}^{\mathrm{T}} \quad \boldsymbol{\theta}_{\mathrm{Tx}}^{\mathrm{T}} \quad \boldsymbol{\varphi}_{\mathrm{Rx}}^{\mathrm{T}} \quad \boldsymbol{\theta}_{\mathrm{Rx}}^{\mathrm{T}}], \tag{2.74}$$

$$\boldsymbol{\alpha} = [\boldsymbol{\alpha}_{\mathrm{HH}}^{\mathrm{T}} \quad \boldsymbol{\alpha}_{\mathrm{HV}}^{\mathrm{T}} \quad \boldsymbol{\alpha}_{\mathrm{VH}}^{\mathrm{T}} \quad \boldsymbol{\alpha}_{\mathrm{VV}}^{\mathrm{T}}], \tag{2.75}$$

$$\boldsymbol{\phi} = [\boldsymbol{\phi}_{\mathrm{HH}}^{\mathrm{T}} \quad \boldsymbol{\phi}_{\mathrm{HV}}^{\mathrm{T}} \quad \boldsymbol{\phi}_{\mathrm{VH}}^{\mathrm{T}} \quad \boldsymbol{\phi}_{\mathrm{VV}}^{\mathrm{T}}], \tag{2.76}$$

where $\{\mathrm{HH}, \mathrm{HV}, \mathrm{VH}, \mathrm{VV}\}$ represent four polarimetric transmissions, e.g., HV means horizontal-to-vertical transmission.

2.2 Dynamic Model

A discrete white noise acceleration model is used to describe the changes of propagation parameters [6], with the assumption that the motion and underlying noise process of different parameters are uncorrelated. The discrete-time state transition equation is expressed as

$$\mathbf{x}_k = \mathbf{F}\mathbf{x}_{k-1} + v_k, \tag{2.77}$$

where v_k is state noise following zero mean normal distribution with the variance matrix $\mathbf{Q}$. The state transition matrix $\mathbf{F}$ is formulated as

$$\mathbf{F} = \begin{bmatrix} \mathbf{I}_5 & \mathbf{I}_5 & 0 & 0 \\ 0 & \mathbf{I}_5 & 0 & 0 \\ 0 & 0 & \mathbf{I}_4 & 0 \\ 0 & 0 & 0 & \mathbf{I}_4 \end{bmatrix}. \tag{2.78}$$

The state vector at snapshot k is

$$x_k = [\mu^{\mathrm{T}} \quad \Delta\overline{\mu}^{\mathrm{T}} \quad \boldsymbol{\alpha}^{\mathrm{T}} \quad \boldsymbol{\phi}^{\mathrm{T}}], \tag{2.79}$$

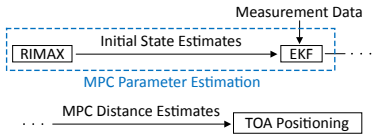


Figure 2.31: Proposed positioning framework.

where the vector $\Delta\overline{\mu}$ contains the velocities of the structural parameters in μ. Here, we intentionally decouple the phase evolution from the MPC tracking to preserve the unique mapping between the phase shift and the distance d_l. The evolution of the state vector from one snapshot to the next is modelled as

$$\begin{aligned}
\mu_k &= \mu_{k-1} + \Delta\overline{\mu}_{k-1} + v_{\mu_k} \\
\Delta\mu_k &= \Delta\overline{\mu}_{k-1} + v_{\Delta\overline{\mu}_k} \\
\boldsymbol{\alpha}_{i,k} &= \boldsymbol{\alpha}_{i,k-1} + v_{\boldsymbol{\alpha}_k} \\
\boldsymbol{\phi}_{i,k} &= \boldsymbol{\phi}_{i,k-1} + v_{\boldsymbol{\phi}_k},
\end{aligned} \tag{2.80}$$

where $v_{[.]}$ denotes the state noise vector. The selection and tuning of process noise variance are very important especially for the narrowband case, because the orthogonality is not tightly held between close-by MPCs. Small variance may lead to smooth but slow tracking, and some small movements might be missed. Large variance enables quick response to non-smooth movements like sharp turns, but with high risk of phase slip. Hence, a trade-off is needed. Here, we follow the guideline that the value of $v_{\Delta\overline{\mu}_k}$ should be in the same order as the maximum acceleration magnitude [6]. The complex path weight is assumed to be slowly varying and to account for larger changes in the propagation processes, e.g., reflection, scattering, etc. Reinitializations of γ_l are sometimes needed in the tracking process [3].

3 Propagation Path Parameters Estimation

As shown in the proposed framework (Fig. 2.31), the MPC parameters are estimated with an EKF. We realize that an accurate initial state estimation is a prerequisite for the fast convergence and accurate tracking in the EKF. In this work, the RIMAX algorithm is applied to the first snapshot for the initial estimates of MPC parameters and noise covariance [4]. Besides, the state dimension adjustment is performed alongside the EKF iteration.

3.1 Initialization with RIMAX

Firstly, MPC detection is performed with a successive path cancellation framework, where the maximum-likelihood (ML)-based 3D-grid approach is used. Each detected path is further optimized locally. This detection framework uses oversampling to enhance delay and angle resolution, therefore closely located MPCs could be detected.

After subtracting the specular-like components, the residual is considered as the colored noise process with a covariance matrix $\mathbf{R}$, which consists of the measurement noise following a Gaussian distribution $\mathcal{N}(0,\ \sigma^2\mathbf{I})$, and the DMC. The DMC is modelled stochastically and the covariance matrix has a shifted Kronecker structure, which is computationally efficient especially when a large antenna array is used [7]. The full noise covariance matrix is given by

$$\mathbf{R} = \mathbf{R}_{\mathrm{R}} \otimes \mathbf{R}_{\mathrm{T}} \otimes \mathbf{R}_f + \sigma^2\mathbf{I}, \tag{2.81}$$

where $\mathbf{R}_f \in \mathbb{C}^{N_f \times N_f}$ is the covariance matrix in the frequency domain with Toeplitz structure. We observed that the power delay profile of the residual shows a spatially white characteristic at the base station (BS) side, therefore the covariance matrices $\mathbf{R}_{\mathrm{T}} \in \mathbb{C}^{N_{\mathrm{Tx}} \times N_{\mathrm{Tx}}}$ and $\mathbf{R}_{\mathrm{R}} \in \mathbb{C}^{N_{\mathrm{Rx}} \times N_{\mathrm{Rx}}}$, which describe the angular distributions at the transmit and receive sides respectively, are assumed as identity matrices in this implementation.

The structural vectors of MPCs and the parameter set of DMC are optimized alternatingly with the Levenberg-Marquardt algorithm and the ML-Gauss-Newton algorithm, respectively. The details can be found in [4].

3.2 Extended Kalman Filter

The path parameters are tracked with an EKF [3]. Due to the non-linear channel model used, we firstly linearize the data model $\mathbf{h}_{sp}(\mathbf{x})$ by taking the first-order partial derivatives over the state vector, which gives the Jacobian matrix as

$$\mathbf{D}(\mathbf{x}) = \frac{\partial \mathbf{h}_{sp}(\mathbf{x})}{\partial \mathbf{x}^{\mathrm{T}}}. \tag{2.82}$$

The first-order and the second-order partial derivative of the log-likelihood function, i.e., the score function $\mathbf{q}$ and the Fisher information matrix $\mathbf{J}$, are also needed in the iteration. These are computed as

$$\mathbf{q}(\mathbf{h}|\mathbf{x}, \mathbf{R}) = 2 \cdot \Re\left\{\mathbf{D}^{\mathrm{H}}(\mathbf{x})\mathbf{R}^{-1}(\mathbf{h} - \mathbf{h}_{sp}(\mathbf{x}))\right\}, \tag{2.83}$$

$$\mathbf{J}(\mathbf{x}, \mathbf{R}) = 2 \cdot \Re\left\{\mathbf{D}^{\mathrm{H}}(\mathbf{x})\mathbf{R}^{-1}\mathbf{D}(\mathbf{x})\right\}. \tag{2.84}$$

The procedure of the EKF is summarized as

$$\hat{\mathbf{x}}_{(k|k-1)} = \mathbf{F}\hat{\mathbf{x}}_{(k-1|k-1)}, \tag{2.85}$$

$$\mathbf{P}_{(k|k-1)} = \mathbf{F}\mathbf{P}_{(k-1|k-1)}\mathbf{F}^{\mathrm{T}} + \mathbf{Q}, \tag{2.86}$$

$$\mathbf{P}_{(k|k)} = (\mathbf{P}^{-1}_{(k|k-1)} + \mathbf{J})^{-1}, \tag{2.87}$$

$$\Delta\hat{\mathbf{x}}_{(k)} = \mathbf{P}_{(k|k)}\mathbf{q}, \tag{2.88}$$

$$\hat{\mathbf{x}}_{(k|k)} = \hat{\mathbf{x}}_{(k|k-1)} + \Delta\hat{\mathbf{x}}_{(k)}, \tag{2.89}$$

where $\mathbf{P}_{(k|k-1)}$ and $\mathbf{P}_{(k|k)}$ are the filter error covariances denoting the prediction and update uncertainties of the state vector, respectively.

3.3 State Dimension Adjustment

In channel sounding, the number of co-existing propagation paths varies over time. The detection and elimination (death-birth) of paths are assumed to be statistically independent and performed alongside the EKF iteration with two separate steps. The first step is to remove unreliable paths by evaluating the relative variance of each path [4], defined as

$$\mathrm{var}_r = \sum_{p=1}^{N_p} \frac{\mathrm{var}_{\gamma_p}}{\mid \gamma_p \mid^2} < \varepsilon_r, \tag{2.90}$$

where γ_p is the magnitude of the estimated path weight of polarization $p \in \{\mathrm{HH}, \mathrm{HV}, \mathrm{VH}, \mathrm{VV}\}$ and var_{γ_p} is the estimation error variance extracted from the filtering error covariance matrix. Intuitively, var_r should be smaller than 0 dB, which indicates that the certainty of the magnitude estimation should be larger than its uncertainty. A reliability check is performed every 30 snapshots and only paths with var_r smaller than the threshold ε_r are preserved in the state for further tracking. Hence, the MPC lifetime is here defined as the time duration that the relative variance of a MPC is below a given threshold, which is geometry-independent in this sense. The next step is to detect new paths. We limit the number of newly initialized MPCs in each snapshot to control the model complexity and reduce the interference between coherent paths [4].

4 Measurement Campaign

This framework is designed for the multiple-input multiple-output (MIMO) case where angular information is available from both sides. However, UEs with single or few antennas are more common in practice and the lack of AODs makes the path estimation and tracking a harder problem. To test the performance of the proposed framework in a real but controlled environment, a measurement campaign was performed in a large sports hall with the RUSK LUND channel sounder. Fig. 2.32 shows an overview of the measurement

Figure 2.32: Overview of the measurement area in the sports hall, Medicon Village, Lund, Sweden. Room dimension is around 20 m × 36 m × 7.5 m.

area. A cylindrical antenna array with 128 ports (Fig. 2.33a) is used as BS at the Rx side, the center of which is 1.42 m above the ground. A conical monopole omnidirectional antenna (Fig. 2.33b) is used to represent a UE at the Tx side. The distance between UE and BS is around 17 m and line-of-sight (LOS) conditions apply. The transfer functions were recorded at a center frequency around 2.7 GHz and with a signal bandwidth of 40 MHz. To avoid large variation of path parameters, especially the phase slip between two consecutive snapshots, the spatial sampling rate of the wireless channel was sufficiently high. In total, 6000 channel snapshots were collected in 19.7 s. The Tx was placed on a tripod and manually moved to write the "**Lund**" letters in a 2 m^3 space. Meanwhile, an optical coordinate measuring machine (CMM) system (Fig. 2.33b) was used to capture the UE motion with accuracy down to millimeter, which acts as the ground truth for performance analysis.

5 MPC Tracking Results and Analysis

This section focuses on the performance of MPC tracking results. Fig. 2.34 shows the tracked propagation distances of MPCs from the EKF implementation. It could be observed that the LOS component with the distance around 17 m is tracked steadily since the beginning. About 2 m apart from the LOS is the ground reflection path which is tracked shortly in the end. Besides, many other MPCs with long lifetimes could also be observed in the range of 20-70 m propagation distance. For better evaluation of the tracking performance, we zoom into the LOS component and compare the distance estimates with the ground truth. The black dashed line denotes the distance estimates from EKF. The

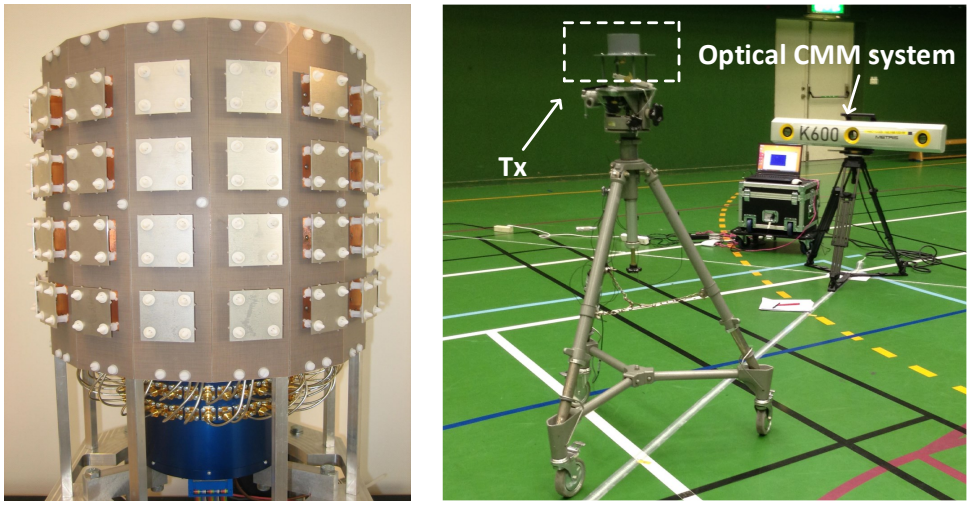


Figure 2.33: (a) The cylindrical antenna array at the Rx side; (b) The conical monopole omnidirectional antenna at the Tx side and the optical CMM system.

red solid line in Fig. 2.35 is the true propagation distance of the LOS component which is calculated based on the 3D coordinates from the optical system and the coordinates of the BS. The two curves are manually time synchronized for better comparison. As shown from the comparison, the EKF could catch all the movements of the UE, even some fine ones and sharp turns. The estimates have a good match with the ground truth most of the time, besides some deviations observed after 16 s. The biggest deviation from the ground truth is about 8 cm. The MPCs located in the same delay bin as the LOS component are correlated and they show degraded quality of tracking.

We further analyzed the angular-power distribution of the tracked MPCs. The MPCs are plotted in a 3D coordinate system based on the estimates of distances and azimuth/elevation AOAs without considering the interaction order. The top view (Fig. 2.36a) shows that the tracked MPCs are distributed over the entire azimuth domain and paths are intensively detected in the similar direction as the LOS component. From the vertical distribution (Fig. 2.36b), a few paths are observed from the ground or at similar height as the BS, but most of the paths are from the complex ceiling structure of the room, e.g., the metal beams of the ceiling in Fig. 2.32. Those complex room structures would bring additional uncertainties to the distance estimates. Moreover, the similar behaviour of the long-tracked MPCs in the angular domain may become a challenge for 3D positioning, for which the MPCs with sparse angles are preferred. However, it is interesting to see the performance in the real but non-ideal case.

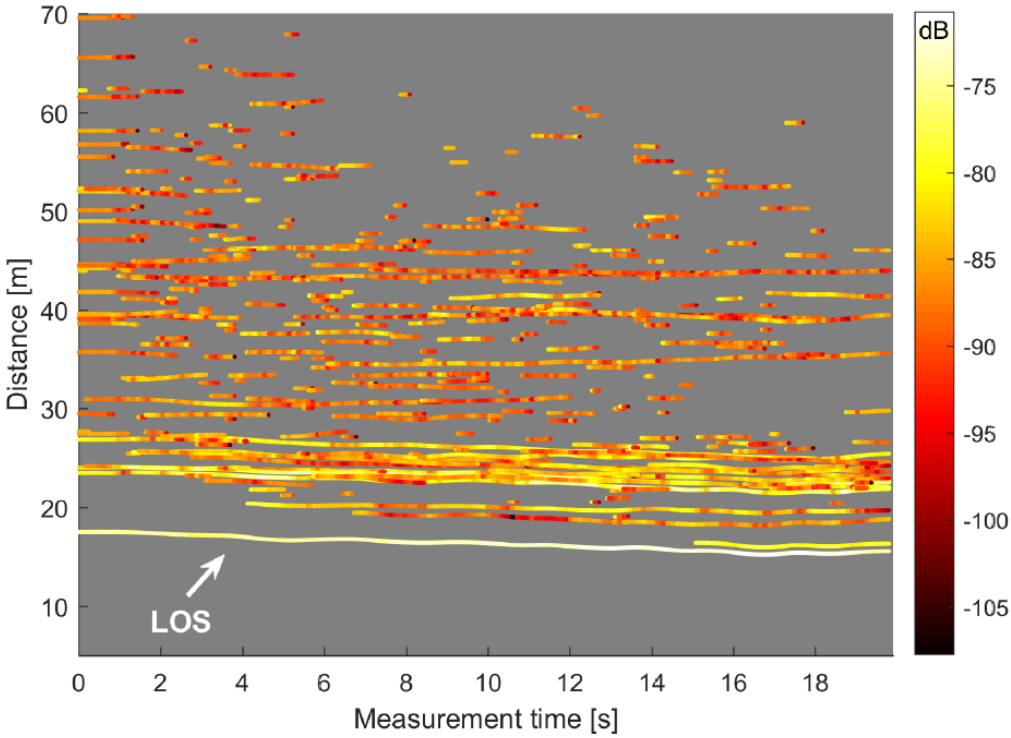


Figure 2.34: The tracked absolute propagation distances of MPCs. The color indicates the power in dB scale.

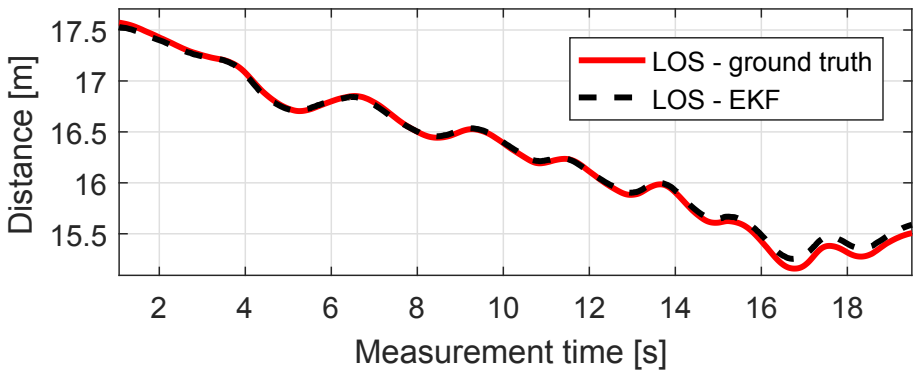


Figure 2.35: Performance evaluation of the tracked LOS component.

Ghost components around some high-power MPCs are observed during the tracking. They usually have similar angles and propagation distances as the dominant MPCs close by and experience very short lifetimes. These components are mainly generated due to power compensation in the estimation procedure and do not have actual physical meaning, therefore

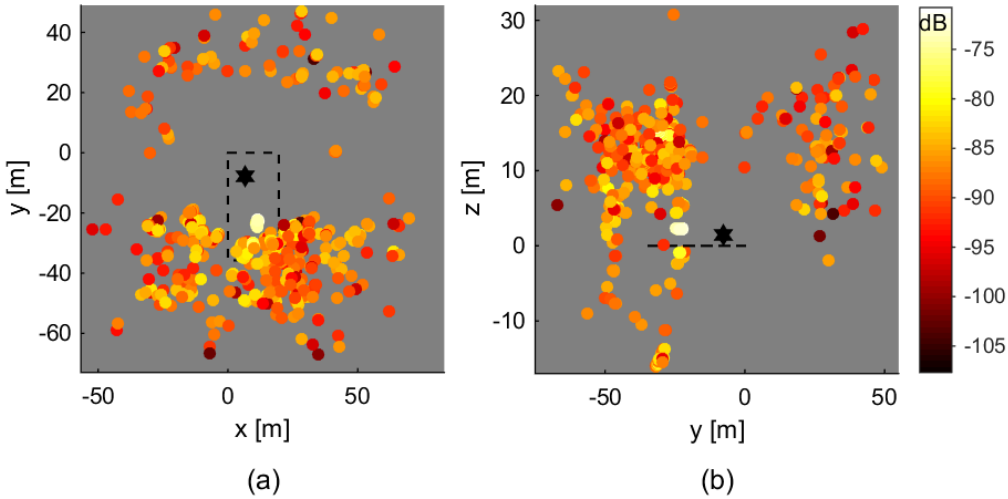


Figure 2.36: 3D plot of the tracked MPCs. Black dashed line denotes the room geometry and the hexagram represents the location of BS. The top-view plot (a) shows how tracked MPCs distributed in the azimuth plane. The side-view plot (b) shows the vertical distribution.

they are not considered in the following positioning step.

6 Positioning Algorithm and Results

As seen in Fig. 2.34 and Fig. 2.35, most of the MPCs can only be observed during fractions of the measurement duration and the estimation quality is not consistent during the whole tracking process for an individual MPC, i.e., there are outliers in the data for which the errors are substantial. Therefore, the question here is how to optimize the positioning performance in the presence of missing data and outliers, which is a highly non-convex problem.

6.1 Mathematical Formulation of Geometrical Problem

A few assumptions are firstly given for modeling the geometrical problem. The tracked MPCs from EKF are numbered with $i = 1, 2, \ldots, m$, where $i = 1$ represents the LOS component. These paths are assumed to originate from n UE positions $\mathbf{Tx}_j \in \mathbb{R}^3$, $j = 1, \ldots, n$. The BS is stationary at position $\mathbf{Rx}_1 \in \mathbb{R}^3$. To formulate the measured distances, we assume that the tracked MPCs are reflected from planar surfaces and mirror the BS, i.e., each MPC can be considered as being received at a mirrored BS position $\mathbf{Rx}_i$. The distance estimates are only given for a set I of (i, j) combinations due to missing data. As it will be shown, the errors of outliers in the distance estimates are substantial. However, for a large

amount of distance estimates the errors are fairly small (in the order of a few centimeters). The measured distances are then

$$d_{ij} = \|\mathbf{R}\mathbf{x}_i - \mathbf{T}\mathbf{x}_j\|_2 + \epsilon_{ij}, \forall (i,j) \in I, \tag{2.91}$$

where $\epsilon_{ij} \in N(0, \sigma_{inl}^2)$ for $(i,j) \in I_{inl}$ and ϵ_{ij} are drawn from an unknown distribution for $(i,j) \in I_{outl}$. This distribution has a significantly larger variance. One useful approach is to minimize the negative log likelihood. To simplify the problem, we assume that the negative log likelihood for the outliers is a constant, i.e., each outlier gives the same penalty. In this way the problem becomes an optimization problem.

Problem 1. *(Time-of-Arrival Self-Calibration) Given absolute distance estimates $d_{ij} \forall (i,j) \in I$, find the inlier set $I_{inl} \subset I$, the UE positions $\mathbf{T}\mathbf{x}_j \in \mathbb{R}^3$ and the mirrored BS positions $\mathbf{R}\mathbf{x}_i \in \mathbb{R}^3$ that solves the following optimization problem*

$$\min_{I_{inl}, \mathbf{R}\mathbf{x}_i, \mathbf{T}\mathbf{x}_j} \sum_{(i,j) \in I_{inl}} (d_{ij} - \|\mathbf{R}\mathbf{x}_i - \mathbf{T}\mathbf{x}_j\|_2)^2 + \sum_{(i,j) \in I_{outl}} C, \tag{2.92}$$

where $I_{outl} = I \setminus I_{inl}$. This is a highly non-linear, non-convex optimization problem. The problem changes character if both $\mathbf{T}\mathbf{x}_j$ and $\mathbf{R}\mathbf{x}_j$ span 3D, or either one of them or both are restricted to a plane or a line as shown in [8]. The problem is ill-defined if there is too little data. For planar problems we require $m \geq 3, n \geq 3$, [9]. For 3D problems more data is needed, typically $m \geq 4, n \geq 6$, [10]. Algorithms for solving Problem 1 using hypothesize and test paradigm are presented in [11].

6.2 Estimation of the Distance Estimates Error Distribution and Mirrored BS Positions

The modified version of Problem 1 where say the transmitter positions $\mathbf{T}\mathbf{x}_j$ are known, is a substantially better conditioned problem. In this case, we can solve for

$$\min_{I_{inl}, \mathbf{R}\mathbf{x}_i} \sum_{i|(i,j) \in I_{inl}} (d_{ij} - \|\mathbf{R}\mathbf{x}_i - \mathbf{T}\mathbf{x}_j\|_2)^2 + \sum_{i|(i,j) \in I_{outl}} C \tag{2.93}$$

independently for each mirrored BS position $\mathbf{R}\mathbf{x}_i$. This can be done by using Random Sample Consensus (RANSAC) [12].

The resulting residuals $d_{ij} - \|\mathbf{R}\mathbf{x}_i - \mathbf{T}\mathbf{x}_j\|_2$ can be used to empirically assess properties of the error distribution. We selected those paths that were longer than 500 snapshots from the tracked $m = 282$ MPCs, which gave a set of 50 MPCs. For each of them, we estimated the mirrored BS position using RANSAC (to obtain I_{inl}) followed by the non-linear optimization of (2.93) (to obtain $\mathbf{R}\mathbf{x}_i$). In total these 50 tracked MPCs gave us

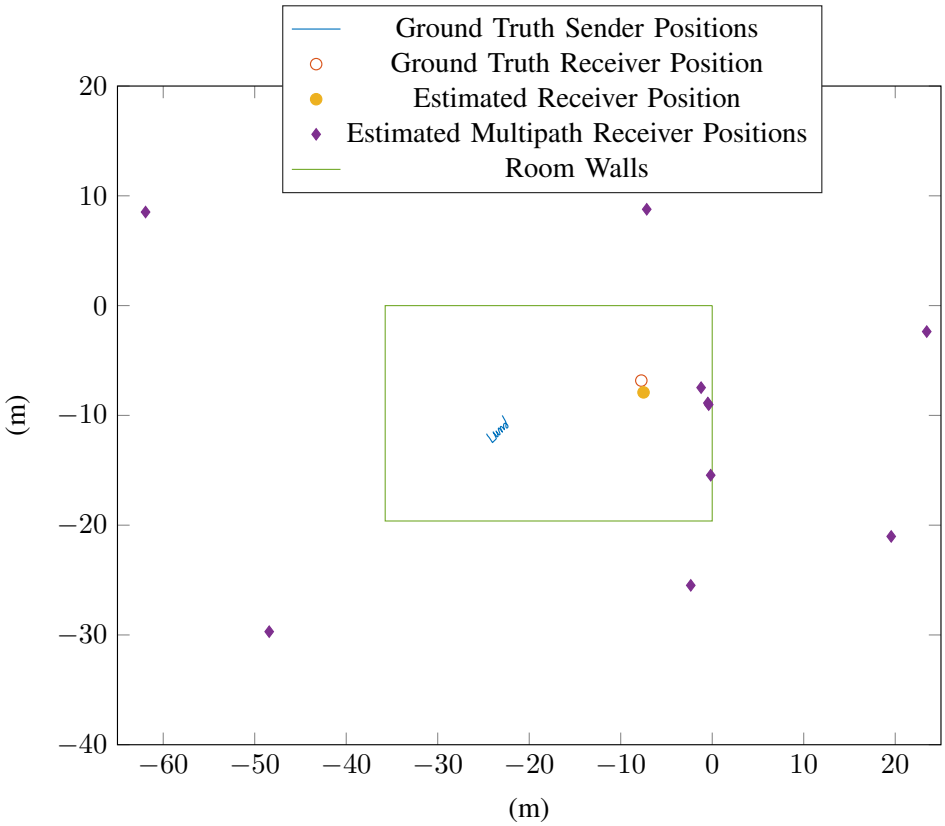


Figure 2.37: Robust estimation of BS position and mirrored BS positions using the ground truth UE positions.

103 480 distance samples, i.e., approximately 2000 each. Of these 77 490 were considered to be inliers. This gives us an estimated inlier ratio of 75%. The standard deviation of the inlier residuals is 4.6 cm.

Using the ground truth UE positions $\mathbf{Tx}_j$, we robustly initialized the LOS component receiver position, i.e., $\mathbf{Rx}_1$ as well as all the mirrored BS positions of MPCs. This was followed by non-linear refinement. The reconstructed BS position $\mathbf{Rx}_1$ along with some examples of the mirrored BS positions are shown in Fig. 2.37. It could be observed that the estimated BS position is located close to the true position, and the mirrored BS positions look plausible.

6.3 Estimation of UE Positions

We now target the full version of Problem 1, where the UE positions, the BS position and all the mirrored BS positions are unknown. To make this highly complex estimation problem tractable, two assumptions are made here. Firstly, $\mathbf{Tx}_j$ are assumed to be constrained in a plane, because the UE was moved approximately in a plane in the measurement. Secondly, we assume that we know which distance estimates are inliers. This problem is then proceeded by splitting the whole dataset in a number of smaller segments in time, which results in 117 segments of length 100 snapshots with 50 snapshots overlap between adjacent segments. For each segment, we robustly initialized both $\mathbf{Rx}_j$ and $\mathbf{Tx}_j$ using minimal solvers and RANSAC [8] based only on the distance estimates from the EKF. This is followed by non-linear optimization. The different solutions from the 117 segments were then registered into a common coordinate system using the overlap between segments. The estimated UE positions (in red) is shown in Fig. 2.38. Also shown (in dashed grey) is the ground truth. The two trajectories have been rigidly registered to each other. It could be observed that the estimated trajectory shows a clear "**Lund**"-word pattern, with all the fine movements details caught. However, the overall shape is stretched along the diagonal direction, which results in a larger deviation from the ground truth especially in the beginning and the end. The largest deviation of the estimated UE position from the ground truth happens at the sharp turn of "**L**", which is 26 cm, and the overall mean deviation is 13 cm. The main reason of the stretch problem is that many MPCs are tracked in a similar direction as the LOS component, as shown in Fig. 2.36. The similar behaviour of MPCs in the angular domain will cause the estimated UE positions to be scaled or projected.

7 Summary and Conclusion

In this paper, we introduced and showed a proof-of-concept for a robust phase-based positioning framework using massive MIMO. MPC parameters, e.g., phase-based distance and angle, are estimated and tracked with an EKF. A TOA self-calibration positioning algorithm is then used for trajectory estimation. The positioning results of a 2D complex movement measurement show that the proposed positioning framework could achieve outstanding positioning performance even with standard cellular bandwidths. Besides, no prior knowledge of the surroundings is needed, so the framework could be easily applied in different environments given that there are enough many scatterers present. To sum up, phase-based positioning using massive MIMO is a promising high-resolution positioning solution for current and next generation cellular systems.

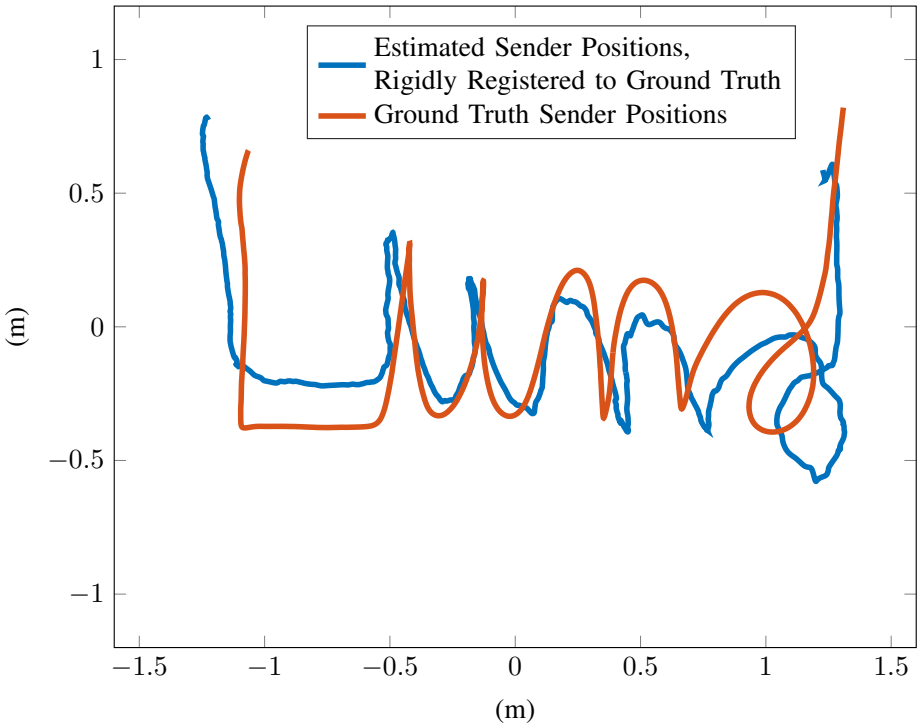


Figure 2.38: The ground truth (dashed-grey) and the estimated UE positions (red). Note that this estimation is only based on the distance estimates and no ground truth UE positions are used.

8 Acknowledgements

The authors would like to thank Erik Leitinger, Jose Flordelis, Joao Vieira and Christian Nelson for helping with the measurements. We would also like to acknowledge the support from Björn Olofsson with the motion capture system. This work was supported by the Swedish Research Council VR and the strategic research area ELLIIT.

Bibliography

[1] E. Leitinger, P. Meissner, M. Lafer, and K. Witrisal, "Simultaneous localization and mapping using multipath channel information," in *2015 IEEE International Conference on Communication Workshop (ICCW)*, pp. 754–760, June 2015.

[2] M. Zhu, J. Vieira, Y. Kuang, K. Åström, A. F. Molisch, and F. Tufvesson, "Tracking and positioning using phase information from estimated multi-path components," in *2015 IEEE International Conference on Communication Workshop (ICCW)*, pp. 712–717, June 2015.

[3] J. Salmi, A. Richter, and V. Koivunen, "Detection and tracking of mimo propagation path parameters using state-space approach," *IEEE Transactions on Signal Processing*, vol. 57, no. 4, pp. 1538–1550, 2009.

[4] A. Richter, *Estimation of Radio Channel Parameters: Models and Algorithms*. PhD thesis, Technischen Universität Ilmenau, Fakultät für Elektrotechnik und Informationstechnik, 2005.

[5] A. F. Molisch, *Wireless Communications*. Wiley Publishing, 2nd ed., 2011.

[6] Y. Bar-Shalom and X.-R. Li, *Estimation with Applications to Tracking and Navigation*. New York, NY, USA: John Wiley & Sons, Inc., 2001.

[7] J. Salmi, *Contributions to measurement-based dynamic MIMO channel modeling and propagation parameter estimation*. PhD thesis, Helsinki University of Technology, Finland, 2009.

[8] S. Burgess, Y. Kuang, and K. Åström, "Toa sensor network self-calibration for receiver and transmitter spaces with difference in dimension," *Signal Processing*, vol. 107, pp. 33–42, 2015.

[9] H. Stewénius, *Gröbner Basis Methods for Minimal Problems in Computer Vision*. PhD thesis, Lund University, Apr. 2005.

[10] Y. Kuang, S. Burgess, A. Torstensson, and K. Åström, "A complete characterization and solution to the microphone position self-calibration problem," in *The 38th International Conference on Acoustics, Speech, and Signal Processing*, 2013.

[11] K. Batstone, M. Oskarsson, and K. Åström, "Robust time-of-arrival self calibration with missing data and outliers," in *2016 24th European Signal Processing Conference (EUSIPCO)*, pp. 2370–2374, Aug 2016.

[12] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, pp. 381–395, June 1981.