TECHNISCHE UNIVERSITÄT
ILMENAU

Doctoral Thesis

Localisation and Tracking of People using Distributed UWB Sensors

Snezhana Jovanoska

Dissertation zur Erlangung des akademischen Grades Doktor-Ingenieur (Dr.-Ing.)

vorgelegt der Fakultät für Elektrotechnik und Informationstechnik
der Technischen Universität Ilmenau

Gutachter: Prof. Dr.-Ing. habil. Reiner S. Thomä
Prof. Dr. Wolfgang Koch
Prof. Dr.-Ing. Dirk Heberling

Vorgelegt am: 17.12.2018
Verteidigt am: 06.02.2020
urn:nbn:de:gbv:ilm1-2020000116

I don't know anything, but I do know that everything is interesting if you go into it deeply enough.

RICHARD FEYNMAN

ABSTRACT

Indoor localisation and tracking of people in non-cooperative manner is important in many surveillance and rescue applications. Ultra wideband (UWB) radar technology is promising for through-wall detection of objects in short to medium distances due to its high temporal resolution and penetration capability. This thesis tackles the problem of localisation of people in indoor scenarios using UWB sensors. It follows the process from measurement acquisition, multiple target detection and range estimation to multiple target localisation and tracking.

Due to the weak reflection of people compared to the rest of the environment, a background subtraction method is initially used for the detection of people. Subsequently, a constant false alarm rate method is applied for detection and range estimation of multiple persons. For multiple target localisation using a single UWB sensor, an association method is developed to assign target range estimates to the correct targets.

In the presence of multiple targets it can happen that targets closer to the sensor induce shadowing over the environment hindering the detection of other targets. A concept for a distributed UWB sensor network is presented aiming at extending the field of view of the system by using several sensors with different fields of view. A real-time operational prototype has been developed taking into consideration sensor cooperation and synchronisation aspects, as well as fusion of the information provided by all sensors.

Sensor data may be erroneous due to sensor bias and time offset. Incorrect measurements and measurement noise influence the accuracy of the estimation results. Additional insight of the targets states can be gained by exploiting temporal information. A multiple person tracking framework is developed based on the probability hypothesis density filter, and the differences in system performance are highlighted with respect to the information provided by the sensors i.e. location information fusion vs range information fusion.

The information that a target should have been detected when it is not due to shadowing induced by other targets is described as dynamic occlusion probability. The dynamic occlusion probability is incorporated into the tracking framework, allowing fewer sensors to be used while improving the tracker performance in the scenario.

The method selection and development has taken into consideration real-time application requirements for unknown scenarios at every step. Each investigated aspect of multiple person localization within the scope of this thesis has been verified using simulations and measurements in a realistic environment using M-sequence UWB sensors.

ZUSAMMENFASSUNG

In vielen Überwachungs- und Rettungsszenarien ist die Lokalisierung und Verfolgung von Personen in Innenräumen auf nichtkooperative Weise erforderlich. Für die Erkennung von Objekten durch Wände in kurzer bis mittlerer Entfernung, ist die Ultrabreitband (UWB) Radartechnologie aufgrund ihrer hohen zeitlichen Auflösung und Durchdringungsfähigkeit Erfolg versprechend. In dieser Arbeit wird ein Prozess vorgestellt, mit dem Personen in Innenräumen mittels UWB-Sensoren lokalisiert werden können. Er umfasst neben der Erfassung von Messdaten, Abstandschätzungen und dem Erkennen von Mehrfachzielen auch deren Ortung und Verfolgung.

Aufgrund der schwachen Reflektion von Personen im Vergleich zum Rest der Umgebung, wird zur Personenerkennung zuerst eine Hintergrundsubtraktionsmethode verwendet. Danach wird eine konstante Falschalarmrate Methode zur Detektion und Abstandschätzung von Personen angewendet. Für Mehrfachziellokalisierung mit einem UWB-Sensor wird eine Assoziationsmethode entwickelt, um die Schätzungen des Zielabstandes den richtigen Zielen zuzuordnen.

In Szenarien mit mehreren Zielen kann es vorkommen, dass ein näher zum Sensor positioniertes Ziel ein anderes abschattet. Ein Konzept für ein verteiltes UWB-Sensornetzwerk wird vorgestellt, in dem sich das Sichtfeld des Systems durch die Verwendung mehrerer Sensoren mit unterschiedlichen Blickfeldern erweitert lässt. Hierbei wurde ein Prototyp entwickelt, der durch Fusion von Sensordaten die Verfolgung von Mehrfachzielen in

Echtzeit ermöglicht. Dabei spielen insbesondere auch Synchronisierungs- und Kooperationsaspekte eine entscheidende Rolle.

Sensordaten können durch Zeitversatz und systematische Fehler gestört sein. Falschmessungen und Rauschen in den Messungen beeinflussen die Genauigkeit der Schätzergebnisse. Weitere Erkenntnisse über die Zielzustände können durch die Nutzung zeitlicher Informationen gewonnen werden. Ein Mehrfachzielverfolgungssystem wird auf der Grundlage des Wahrscheinlichkeitshypothesenfilters (Probability Hypothesis Density Filter) entwickelt, und die Unterschiede in der Systemleistung werden bezuglich der von den Sensoren ausgegebene Informationen, d.h. die Fusion von Ortungsinformationen und die Fusion von Abstandsinformationen, untersucht.

Die Information, dass ein Ziel detektiert werden sollte, wenn es aufgrund von Abschattungen durch andere Ziele im Szenario nicht erkannt wurde, wird als dynamische Verdeckungswahrscheinlichkeit beschrieben. Die dynamische Verdeckungswahrscheinlichkeit wird in das Verfolgungssystem integriert, wodurch weniger Sensoren verwendet werden können, während gleichzeitig die Performanz des Schätzers in diesem Szenario verbessert wird.

Bei der Methodenauswahl und -entwicklung wurde die Anforderung einer Echtzeitanwendung bei unbekannten Szenarien berücksichtigt. Jeder untersuchte Aspekt der Mehrpersonenlokalisierung wurde im Rahmen dieser Arbeit mit Hilfe von Simulationen und Messungen in einer realistischen Umgebung mit UWB Sensoren verifiziert.

ACKNOWLEDGEMENTS

The work for this thesis was carried out in part at the Electronic Measurements Research Lab at the Technische Universität Ilmenau, Ilmenau, Germany and in part at the Sensor Data and Information Fusion department at the Fraunhofer Insitute for Communications Information Processing and Ergonomics (FKIE), Wachtberg, Germany. I would like to take this opportunity to thank all the people who helped me complete this work.

First and foremost I would like to thank my supervisors, Prof. Reiner Thomä from TU Ilmenau and Prof. Wolfgang Koch from Fraunhofer FKIE for their continued support and guidance through the years. Their many ideas and suggestions have helped in defining my thesis topic as a bridge between electronic measurements and statistical fusion and tracking theory.

Furthermore, I would like to thank Prof. Dirk Heberling for accepting to review my work, Prof. Jochen Seitz for chairing the thesis committee, as well as Prof. Matthias Hein and Prof. Giovanni Del Galdo for participating in my thesis committee.

I would also like to thank all my former and current colleagues at the EMS research group from TU Ilmenau and at the SDF department at Fraunhofer FKIE for providing a pleasant work environment, continuous collaboration and for the many interesting scientific and non-scientific discussions. In particular, I would like to thank Dr. Rudolf Zetik for all the UWB related discussions and his help in conducting the measurement campaigns, Dr. Felix

Govaers for his invaluable advice and support while working on my thesis and Dr. Martina Brötje for her willingness to openly share her extensive knowledge in multistatic fusion and tracking. I would also like to thank the late Mr. Klaus Wild for his early-on invested interest in my research work and for integrating me in the research groups as Fraunhofer FKIE in 2015, where I could work on a more hands on research and many different application scenarios and demos. They have broaden my focus and general research interests.

A special thanks goes to my boyfriend Martin Käske for his unconditional support and all the time he spent discussing probability theory and measurement setups with me.

Last but not least, I would like to thank my family and friends for their constant support and encouragement, and for reminding me of the important things in life on every turn.

Мама и тато, без вашата поддршка и охрабрување оваа работа не би била возможна. Ви благодарам за сè!

To my parents

CONTENTS

LIST OF FIGURES

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Appendix III

Appendix IV

ABBREVIATIONS

2D two-dimensional

3D three-dimensional

AOA angle of arrival

CFAR constant false alarm rate

CRLB Cramér-Rao lower bound

DC direct current

DSP digital signal processing

EKF Extended Kalman filter (KF)

EM electromagnetic

EWMA exponential weighted moving average

FCC Federal Communications Commission

FISST finite set theory

GM Gaussian mixture

GMPHD Gaussian mixture probability hypothesis density (PHD)

GPS global positioning system

IPCP inter-period correlation processing
IRF impulse response function

JDL Joint Directors of Laboratories

KF Kalman filter

LIDAR Light Detection and Ranging
LOS line-of-sight
LTI linear time invariant

MAP maximum-a-posterior
MHT multiple hypothesis tracking
MIMO multiple input multiple output
MMSE minimum mean squared error
M-sequence maximum length binary sequence

NLOS non-line-of-sight

OSPA optimal subpattern assignment

pdf probability density function
PGF probability generating function
PGFL probability generating functional
PHD probability hypothesis density
PL path loss

RCS radar cross-section

RF radio frequency

RFS random finite set

RSS received signal strength

SMC sequential Monte Carlo

SNR signal-to-noise ratio

TDOA time difference of arrival

TOA time of arrival

UKF Unscented KF

UWB ultra wideband

WLAN wireless local area network

LIST OF SYMBOLS

σ_{RCS} target radar cross section

c_0 signal propagation velocity in air

τ Time delay

β effective bandwidth
B_{frac} fractional bandwidth
B absolute bandwidth

T_p pulse width

f_U upper frequency
f_L lower frequency
f_C center frequency

a_{Tx} stimulus signal

b_{Rx} receive signal

δ Dirac delta function

Λ_{21} scatterer response for incidence by transmitter (1) and observation by receiver (2)

η_{Rx} receive antenna efficiency

η_{Tx} transmit antenna efficiency

G_{Rx} receive antenna gain

G_{Tx} transmit antenna gain

P_{Rx} receive power

P_{Tx} transmit power

d_{Tx} range distance from transmitter to target

d_{Rx} range distance from target to receiver

$d_{Tx,Rx}$ distance between transmitter and receiver

γ_{τ} minimum resolvable delay

D_{min} minimum detectable distance

$r_{s,j}$ true range of a target with respect to sensor s and receiver j of sensor s

$\tilde{r}$ estimated range of a target

ε distance (range) estimation error i.e. additive noise of TOA estimation

A area of interest

T_s intersection threshold for sensor s

p probability function

$\psi_{z,t}$ multiple target likelihood function

$\mathcal{N}$ Gaussian density function: $\mathcal{N}(:, m, P)$ with mean m and covariance P

$\mathbf{x}_t$ state vector of a target at time t with coordinates (x_t, y_t) in two-dimensional (2D) and velocity $\dot{x}_t$ and $\dot{y}_t$ in the respective dimension

P covariance matrix

X_t set of target states

Ξ random finite set of target states

$N_{\mathbf{x}_t}$ number of targets at time t

$Z_{1:t}$ set of received observations/measurements from start up to time t

$\mathbf{z}_t$ target observation at time t

Z_t set of observations

Ψ random finite set of observations

$N_{\mathbf{z}_t}$ number of observations at time t

S random finite set of surviving targets from previous time step

C random finite set of clutter

κ_t clutter intensity

Θ random finite set modelling measurements generated by targets

Γ random finite set of newly born targets

γ_t spontaneous birth intensity
$v_{S,t|t-1}$ survival intensity

N_s number of sensors
N_{sj} number of receivers of sensor s

v_t posterior intensity function
$v_{t|t-1}$ propagated intensity function
$v_{D,t}$ intensity of the detected targets

J_{γ_t} number of Gaussian mixture (GM) components used for target birth

J number of GM components

τ pruning threshold

U merging threshold

T merging threshold

$\tilde{\mathbf{x}}_t$ particle sample of a distribution

N_p number of particles

w weight of a GM component

m mean of a GM component

σ standard deviation of a GM component

p_D probability of detection

ϱ_F false alarm intensity

p_S probability of survival

υ_t process noise
Q_t process covariance
ω_t observation process noise
R_t observation process covariance

ζ generic target state

z generic target observation

X generic target state set

$\phi_{t|t-1}$ transition density

$\beta_{t|t-1}$ spawned target birth intensity

F_t transition matrix

H_t observation model matrix
$h_{s,j}$ observation model function or measurement equation

K Kalman filter gain

q_t clutter/false alarm probability distribution over the measurement space

λ_t Poisson parameter defining the expected number of false alarm

s smooth transition function

θ clockwise angular extent of a target

ϕ counterclockwise angular extent of a target

CHAPTER 1

INTRODUCTION

Detection and localisation of people in indoor environments is of great interest in many applications. From detection of first responders in a burning or collapsing building, detection of users in a smart home environment for improvement of the user experience to many security and surveillance applications, people detection and localisation in known or unknown indoor scenarios opens up the road for improvement of current and emergence of new methodologies and applications.

When talking about indoor localisation, it is often meant localisation of collaborative objects. This is often referred to as active localisation since the object to be localised collaborates with the localisation system, for example by carrying radio frequency (RF) identification chips. Localisation and tracking of non-cooperative objects is often referred to as passive localisation. Passive localisation also refers to localisation of objects using passive systems such as cameras or RF systems in which the system itself does not emit any signals. To make the distinction, passive localisation where the objects to be localised do not cooperate with the localisation system is also referred to as non-cooperative or device free localisation [1, 15–17]. A review of localisation schemes and systems is given in Appendix I.

Indoor localisation and tracking of people in passive manner has many applications such as intruder detection, emergency response, surveillance and security, elderly care, smart homes, etc [18–21]. Different techniques for

person detection and tracking are developed, such as vision-based systems, Light Detection and Ranging (LIDAR), infra-red and radar. LIDAR and vision systems fail when visibility conditions are poor as in smoky, foggy or dusty environments, whereas radar is resilient to adverse environmental influences such as visibility conditions. Narrow-band RF-based solutions are presented in [22, 23], where the body influenced shadowing over the transmission link is exploited. More recent works [24, 25] provide improvement by partitioning the environment into cells. Narrow-band RF-based solutions are low-cost and attractive for indoor localisation, however they are susceptible to multipath fading, making it difficult to create models for dense clutter environments. Ultra wideband (UWB) radar is applicable for short to medium distance target detection and localisation due to its high spatial and temporal resolution and inexpensive circuitry [26–28]. Compared to visual sensors, UWB sensors can maintain their accuracy even in bad visibility conditions. UWB is promising for indoor positioning due to its high resolution ranging and obstacle penetration capabilities [29–32].

UWB signals have a very good time resolution due to the large bandwidth and allow for centimetre accuracy in ranging. Thus they can detect objects close to each other and even multiple echoes per object. The low frequencies of the UWB spectrum can be used for penetration through objects, allowing imaging, localisation and/or tracking of obstructed targets. UWB systems preserve their advantages such as high resolution and robust operation even in multipath rich environments [33]. Thus UWB systems can operate regardless of the visibility conditions (dark or smoke-filled environments) and through non-metal obstacles (e.g. through-wall) [34]. The UWB technology can be used in various applications in diverse fields, such as medical technology [35], through-wall imaging [36], mining (for detection of hidden or damaged structures) [37], civil engineering and building construction [38], rescue of trapped people [39], detection of land-mines [40], autonomous navigation in buildings [41] etc. Within this thesis detection and localisation of multiple people in indoor scenarios is considered. The envisaged applications include security applications i.e. assistance to police by pro-

viding information regarding the number and activity of people in a room or building before entering (use of through-wall detection capability) and assistance to fire fighters by proving them information on the whereabouts of possibly injured people in smoke-filled areas. Other possible applications of the system can be found in assisted living and intelligent homes, however these applications have not been directly taken into consideration for the work presented here.

Despite the high range resolution capabilities of the UWB radar, detection and localisation performance can be significantly improved by cooperation between spatially distributed nodes forming a sensor network. UWB sensors have limited coverage due to the low power used as well as the interference of coexisting systems. Shadowing induced by objects in the area of interest introduces additional limitations to the sensor coverage area. Distributed sensor nodes can also acquire comprehensive information for the structure of an unknown environment and construct an electromagnetic image relative to the sensor-to-sensor coordinate system. Distributed observation allows robust detection and localisation of passive objects and identification of certain features of objects such as shape, dynamic parameters and time variant behaviour. When using multiple distributed UWB sensors the coverage area of the system is increased and the location estimates of detected persons are refined by using multiple detections of the same person by multiple sensors. Additionally, objects that can not be detected by one of the network sensors, would be detected by another sensor. The distributed UWB sensor nodes can function using different sensory properties, such as only transmitter or only receiver or as an autonomous operating sensor containing a complete multistatic UWB radar platform consisting of transmitter and receivers [26, 32]. The cooperation of the sensor nodes in the sensor network, as well as between the sensors and the fusion platform needs to be taken into consideration. Another aspect to be taken into consideration is how the information from the multiple sensors can be fused to enhance the desired information. A likelihood function denotes the information provided by the sensors and provides a score based on the quality and consistency of this information.

By using distributed sensors the area of interest can be better observed and shadowing influences are reduced by being able to observe an area of the scenario shadowed to one sensor by another sensor. Another way to deal with the shadowing problems is to analyse possible target missdetections and if caused due to shadowing, use this additional information within the detection and tracking system. When a sensor expects to detect a target within its field of view, if this target is not detected, there are three possible causes: sensor failure, target is out of range or target is occluded. If the sensor is functioning and there is reasonable confidence that the target is still in the sensor field of view, it can be concluded that a target is occluded. This information can now be used within the system and add additional insight for the presence and location of targets within the area of interest. In this thesis negative information fusion is considered and a concept for its incorporation within the existing framework is suggested.

If the available information regarding the location is insufficient or noisy, tracking can be used to improve location estimates. Tracking as a method for estimation of the current state of a system by using the information of the immediate previous state of the same system and information of the propagation of the changes in the system has been widely used for estimation of target locations. In addition to the noisy observations, position estimates based on the previous states of the detected people are used to improve the current location estimates. The Kalman filter (KF) is the optimal solution for single target Bayes filter in linear Gaussian systems [42, 43]. Its variants Extended KF (EKF) and Unscented KF (UKF) are wide-spread for non-linear problems [42, 44–47]. Extensions for multiple state tracking include parallel KF banks, which require observation-to-track association before tracking is performed (also known as Gaussian sums filters) [48]. For multiple target tracking different methods have been developed over the years. Many require the number of targets to be known and use measurement-to-track association before estimating the target states [49–52]. These kinds of methods require large memory, have high computational complexity and require data association which is a problem when targets

are close to each other or when there is a large number of observations.

In unknown environments, one of the main requirements for a good multiple target tracking approach is the possibility to estimate the unknown varying number of targets as well as their states from noisy and incomplete observations over time. Problems arise already with the need to jointly estimate the number of targets present and their states. Other problems include the random unknown multiple number of targets, observations per target, the detection uncertainty of the system, multiple unknown clutter detections, etc. A combinatorially and computationally less complex solution that does not require explicit data association is the probability hypothesis density (PHD) filter [53–55]. It works on the principle filter before track. It treats all targets, or objects of interest, as a set valued state and all available observations as a set valued observation. It is based on finite set statistics [53]. It does not require the number of targets to be known, since it jointly estimates both the number of targets and their state. It has been shown that it can operate in environments with false alarms and missed detections [53]. Since the PHD filter does not require data association and is not keeping various hypothesis of possible measurement to track assignments, its combinatorial and computational complexity, as well as memory requirements, are lower compared to other multiple target tracking methods, such as the multiple hypothesis tracking (MHT) algorithms.

In recent research works the derivation of PHD intensity filters based on point processes has been presented as a generalisation to the derivation via random finite sets (RFSs). Within the framework probability generating functions (PGFs) and probability generating functionals (PGFLs) are used to derive the filter equations. This framework is especially appropriate for generalised multi-target multi-sensor problems [56]. Since the work presented in this thesis focuses on the practical application of multiple sensor multiple target tracking filters, the derivation of the filter equations and their comparison is not discussed in detail. More on the derivations can be found in [57]. In this thesis the PHD filter as implemented by Vo et al. in [54] using Gaussian mixture (GM) approximation is used. Due to

the restrictions of the PHD filter in its derivation and assumptions taken, extensions have been made to adapt it for the specific real world scenarios used within this thesis, including multiple target detections per sensor.

Within this work an indoor localisation system for multiple targets is developed using maximum length binary sequence (M-sequence) UWB sensors. The main requirement set on this system is the possibility to detect and localise multiple targets in real time in possibly unknown scenarios. At first a simple single sensor solution is considered, followed by an extension using a distributed sensor network constellation. System improvements are provided by the use of the PHD filter for data fusion and tracking. Sensor missdetections and dynamic shadowing models are exploited for improved target tracking when the number of sensors is restricted. An overview of the thesis contributions is given in the following section.

1.1 Thesis contributions and proposed solutions

UWB device-free localisation is able to overcome many of the drawbacks of narrowband signals, however it still has a number of unresolved issues. In this thesis indoor localisation and tracking of multiple people using UWB sensors is considered. UWB sensors are used for people detection using signals reflected off their bodies. In case of multiple targets, it can often happen that targets further from a sensor are often not detected. The intended application scenarios are unknown scenarios for which apriori data can not be gathered and used for improvement of the target detection and tracking methods.

The following contributions are presented in this thesis:

↪ A comprehensive framework for real-time person detection and tracking using UWB sensors is presented. All methods needed for processing received signals are presented with different alternatives which can be used on scenario to scenario basis and detailed description of the selected methods for multiple person detection and localisation in real-time.

↪ Multiple person detection and range estimation approach for real-time application is presented. To extract the target information, first the signals impulse response function (IRF) is aligned to 'time zero' using the direct signal path from the transmitter to the receiver and their physical distance. Then a background subtraction technique which can work online and in unknown environments is applied to remove or reduce the influences of all static objects in the environment and the direct path contribution. Due to the extent of a person in the UWB signal, a multiple target range estimation method which takes this into consideration is used. The method used extends the constant false alarm rate (CFAR) estimation using clustering techniques. Range tracking has also been considered for clutter suppression.

↪ A multiple target localisation approach is proposed where an association method is proposed to assign target range estimates to the correct targets.

↪ The work in this thesis is directed toward the practical application of the presented methods. Thus practical challenges for multiple target detection and localisation are discussed throughout this thesis. The target missdetection problem in multiple target scenarios is explored for which two solutions are proposed: using multiple sensors or incorporating negative information within the framework.

↪ A concept for distributed UWB sensor network with real-time functionality is presented. The intended scenario for application is an unknown scenario where the sensor network can be quickly set up and directly used. As a rule of thumb it is advised that the sensors are placed around the room or building which should be inspected with as much separation as possible, parallel to the walls (if the area of interest is not encapsulated by walls, direct line of sight to possible targets should be preferred to reduce multipath influences), and whenever possible the sensors of the network should have *cross view* of the area of interest, i.e. each sensor node should be looking at the scenario

from a different angle. Two modes of sensor operation have been explored: sequential transmitter operation which avoids interference and parallel transmitter operation where each sensor is active at all times. Due to hardware restrictions, and insignificant transmitter interference, parallel operation is considered in the work presented in this thesis and the experiments conducted. Sensor cooperation and synchronisation are also taken into account, as well as fusion of the information provided by all sensors.

↪ Multiple target tracking framework based on the PHD filter for bistatic range-only target observations is proposed. Both centralised and distributed fusion approaches are explored as solutions for multiple sensor fusion and tracking. For the first method, single sensor location estimates from all detecting sensors are used as target observations within the tracking framework. The second method considers each transmitter-receiver pair in the sensor network as a separate sensor. The range estimates are then directly used as observations in the multiple target tracking framework based on the PHD filter with EKF measurement equations.

↪ Lastly, a framework for exploiting the missdetections of the targets by the sensors is developed. By characterising the reason why a target may not be detected, additional information on the target state may be utilised. This methodology is also known as negative information fusion. By characterising the occlusion regions a detected target may have over the scenario, a conclusion may be drawn that if a target is not detected, it is likely that it is currently in one of the occluded regions of the scenario. In this thesis the probability that a target is occluded by another target from the scenario is incorporated within the target tracking framework.

All derived methods and frameworks have been experimentally verified for real-time application on numerous online and offline demonstrations.

1.2 Own publications

In [2], an overall method starting from measurement set-up and signal acquisition to multiple sensor multiple person tracking is defined. Each sensor processes the received impulse responses and localises any detected people. The location estimates from each sensor are then transmitted to a fusion center where they are fused to represent the full view of the environment and all moving objects within it. This method is also termed location information fusion or distributed fusion approach within this thesis. [3] is an elaborate chapter written in collaboration with many other researcher describing cooperative localisation and object recognition in autonomous UWB sensor networks. It includes a summary of the work presented in [2], as well as active node localisation as discussed in [4, 5].

The centralised fusion approach is considered in [1]. Each sensor transmits the range estimates of any detected target in each transmitter-receiver channel. Multiple sensor likelihood function is derived by using all range estimates. An experimental demonstration framework is defined for real-time application and discussed in a scientific report in [6]. The advantages of using multiple distributed sensors is demonstrated and discussed in [7]. The full multiple target detection and tracking framework and comparison of the centralised and distributed method is presented in [8].

In [9, 10] dynamic shadowing induced by people in the scenario within the localisation and tracking framework is considered. A model for the dynamically shadowed region by a target is derived and the occlusion likelihood function is integrated in the multiple target tracking framework. By using the negative information the state of the area of interest and the people within it can be observed by less sensors.

Other publications with my contribution which are not discussed in detail in this thesis include [11] where a simple method for imaging based multiple person localisation is presented, [4, 5] where active node localisation in realistic indoor environments (multipath and non-line-of-sight (NLOS)) is presented, [12] where stationary object detection in unknown environments

is discussed and [13] which discusses detection and localisation of a single target located behind a corner from the sensor by exploiting the multipath components and reflections off the walls in the scenario. In [14] a concept for heterogeneous multiple sensor fusion system is discussed, where radar, cameras and RF direction finder are used for detection and localisation of small commercial drones. This publication is also not covered in this thesis since it is out of scope.

1.3 Thesis outline

The thesis is structured as follows:

- ↪ Chapter 2 describes the target detection and localisation principle using an UWB system. It first describes the UWB sensing principle with electromagnetic (EM) wave propagation in the scenario and definition of an IRF. A bat-type UWB sensor node structure is described as used in the experiments done for verification of the methods described in this thesis. The person detection principle and methods used are described in Section 2.5. Target echo separation techniques and range estimation for multiple targets are described as applied in the system framework. The person localisation procedure for single targets and a derived extension for multiple targets based on the sensor geometry is described in Section 2.6.
- ↪ Chapter 3 describes the conceptual design and architecture of a UWB sensor network. Details on the sensor network possibilities for operation and the sensor node and fusion center frameworks as used in the network architecture are given. Distributed sensor fusion with its challenges are also described in this chapter. Multiple sensor likelihood function is derived and evaluated for both centralised and distributed multiple sensor fusion approaches.
- ↪ Chapter 4 starts by giving some background on Bayesian tracking methodology for single and multiple targets. Typically used approaches are shortly described with more details given on applied

methods (PHD filter). The two approaches, location tracking and direct range-to-location tracking are described, compared and experimentally evaluated in an office scenario.

↪ Chapter 5 describes an extension to the multiple target tracking framework, where in addition to the sensor observations, non-detections of the targets by the sensors is also considered (also known as negative information fusion). The region in the scenario which is being occluded by a moving target is first defined and then derived as an occlusion likelihood function. The occlusion likelihood is integrated within the framework and the information gain is evaluated in a simulation as well as measurement scenario.

↪ Chapter 6 summarises the presented work and gives some perspectives for future research.

CHAPTER 2

ULTRA WIDEBAND-BASED LOCALIZATION

Person localization based on UWB can be described as a two step approach. The first step is estimating parameters from the IRF, such as time delay based parameters, Doppler, angle, signal strength, etc. The second step is using the estimated parameters to derive the location of the person. Different localization schemes based on the available parameters are presented in Appendix I. Although two-step approaches are suboptimal, they have much lower complexity compared to direct approaches where the position is estimated directly from the signals travelling between the nodes and the performance of the two approaches is close for sufficiently high signal-to-noise ratio (SNR) and signal bandwidth [58, 59].

This chapter introduces the working principle of UWB sensing. It also defines a UWB sensor node as an autonomous unit comprised of antennas, RF components and digital signal processing (DSP) unit as used later in this thesis. For verification of the methods described in this thesis UWB sensors designed by Ilmsens GmbH (formerly part of TU Ilmenau) [60] are used. The person detection principle and time of arrival (TOA) estimation techniques are described in Section 2.5 of this chapter. The principle of TOA based person localisation using a single UWB sensor node for single and multiple targets is described in Section 2.6.

2.1 Introduction

UWB technology has been found useful in many industrial, medical, surveillance, security, search and rescue applications among others. Whereas high-power medium and long range radar systems are reserved for military use, high resolution, short range and low power (<1 mW) devices are of interest in civilian applications. The Federal Communications Commission (FCC) defines UWB signals as signals with fractional bandwidth greater than 20 % (if the center frequency is below 2.5 GHz) or absolute bandwidth of at least 500 MHz (if the center frequency is above 2.5 GHz). The fractional bandwidth is defined as $B_{frac} = \frac{B}{f_C}$, where f_C is the center frequency i.e. $f_C = \frac{f_U + f_L}{2}$. The absolute bandwidth is calculated as the difference between the upper, f_U, and the lower, f_L, frequency of the -10 dB emission point, i.e. $B = f_U - f_L$. Due to the large bandwidth, UWB systems are characterised by short duration waveforms. A good overview on UWB technology, sensing, radar architecture, and some applications is given in [61].

UWB signals are non-ionising (compared to X-rays) and require no contact with subjects, thus are considered safe to humans [34, 35, 62, 63]. UWB signals occupy a much wider frequency band compared to narrowband signals and thus share the spectrum with other existing systems. Due to the low emission levels allowed by regulatory agencies , UWB systems are mainly used for short-range indoor applications. UWB systems are ideal for use in frequency sensitive places such as hospitals due to the low power and high precision capabilities. EM waves stimulate the environment and dielectric changes in the propagation medium induce reflections, scattering or reflections of the EM waves which are received and analysed by the UWB sensors. Based on the recorded echoes different properties of the inspected environment may be identified, such as different geometrical structures, presence and position of people in the environment, possible hidden weapons, etc.

Since this thesis is concerned with person detection and localisation, in addition to presenting the general UWB sensing principle, in this chapter

person detection and localisation is discussed, taking into consideration scenarios with multiple people. A simple extension to the single person localisation method is proposed for multiple person localisation.

2.2 UWB sensing

UWB sensors use low power EM waves for sensing the environment as shown in Fig. 2.1. The objects in the scenario influence the wave propagation. The received signal is then distorted based on the object geometry and material properties. Based on the antenna characteristics, the observation volume is either cone-shaped or omnidirectional. In the time domain radar equation, all objects that do not move too fast with respect to the hardware such as antennas and scatterers may be considered as an linear time invariant (LTI) systems. The information about the channel propagation (transmission, reception and scattering) can be formally described by IRFs.

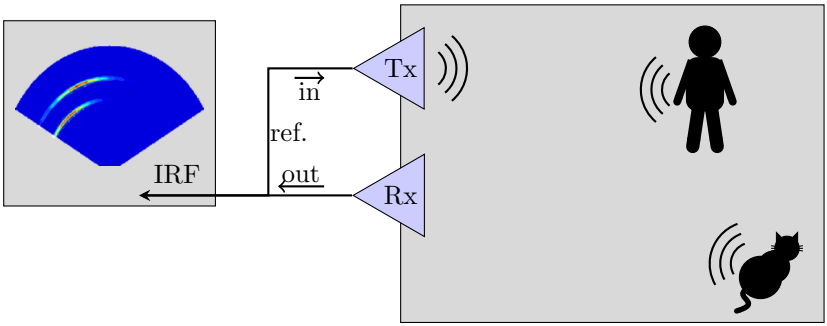


Figure 2.1: Two target scenario and corresponding impulse response function obtained by a bistatic radar

Let the transmitting antenna be denoted by Tx and the receiving antenna be denoted by Rx. Disregarding polarization, dispersion, angular and range dependencies, and assuming a limited target size in the far field and no antenna coupling, the transmission between the two antennas can be written

in terms of M scatterers:

$$b_{Rx}(t) = S_{21}(t) * a_{Tx}(t) \;\Rightarrow\; b_{Rx}(t) \sim S_{21}(t) \text{ for } a_{Tx}(t) \approx \delta(t)$$
$$\text{and } S_{21}(t) = T_1(t) * R_2(t) * \sum_{i=1}^{M} \frac{{\Lambda_{21}}^{(i)}\left(t - \frac{{d_{Tx}}^{(i)} + {d_{Rx}}^{i}}{c_0}\right)}{{d_{Tx}}^{i}\,{d_{Rx}}^{i}} \tag{2.1}$$

where a_{Tx} is the stimulus signal, b_{Rx} is the received signal, $c_0 = 3 \times 10^8$ m/s is the signal propagation velocity in air, ${d_{Tx}}^{i}$ and ${d_{Rx}}^{i}$ are the distances of the i-th scatterer to the transmitter and receiver, T_1 and R_2 are the antenna responses for transmit and receiving mode, and ${\Lambda_{21}}^{i}$ is the scatterer impulse response for incidence by Tx and observation by Rx and contains all information about the i-th target accessible by the measurement. One thing that is not mentioned in the above equation is the antenna coupling and the object interaction which also constitutes part of the IRF. Such interactions include multipath components. ${\Lambda_{21}}^{i}$ responses can be interpreted either as a reaction of a single body onto an incident field or the reaction of distinct scattering centres of a composed target. The scattering response $\Lambda_{21}(t)$ of a complex structured target is typically comprised of many peaks caused by specular reflections and damped oscillations that represent the eigenmodes of the target. In this equation clutter is considered as some of the M scatterers. To resolve these properties the temporal width of the sounding wave must be shorter than the peak distance and the sounding bandwidth should cover the eigenfrequencies of the target [61].

The stimulus signal a_{Tx} may have different waveforms, e.g. chirp signal, short-impulse signal, binary sequence, etc. An appropriate stimulus signal should be used depending on the application where it is required. Chirp signals are not appropriate for real-time surveillance due to their slow measurement rate. Short-impulse signals also result in low measurement rate and these systems are susceptible to jitter and drift. Pseudo-random binary sequences allow for real-time operation, are generated in stable manner and have low crest factor (allowing signal energy maximisation at low peak voltages).

The UWB sensing system used for the work in this thesis is an M-sequence UWB radar [64] which uses a multiple length binary sequence as stimulus signal. The M-sequence generator can be realised on integrated circuit technology, meeting the bandwidth, low jitter and high SNR requirements [40, 61, 64]. Additionally, its spectral shape follows the $(\text{sinc})^2$ function with around 80 % of its energy concentrated at frequencies below half of the clock rate [61].

Since the stimulus sequence is cyclic, an exact 'time zero' for the IRF can not be directly determined. To determine 'time zero' calibration measurements should be performed.

Using a UWB sensor, the IRF $S_{21}(t)$ is primarily determined from the inspected environment. The IRF contains information about possible targets and the environment based on their reflections. In the applications considered in this thesis, a non-cooperative target is assumed. The EM waves scattered from its body are used to aid in the detection of the target. The IRF is determined by cross-correlating the excitation signal with the received signal.

Rewriting Eq. (2.1) in terms of the auto- (C_{aa}) and cross-correlation (C_{ba}) function of the stimulus and receive signal, the following equation is obtained:

$$\begin{aligned} C_{ba}(t) &= S_{21}(t) * C_{aa}(t) \\ \text{with } C_{ba}(t) &= b_{Rx}(t) * a_{Tx}(-t). \end{aligned} \tag{2.2}$$

For $C_{aa} \approx \delta(t)$, $C_{ba} \propto S_{21}(t)$, i.e. the desired IRF $S_{21}(t)$ is proportional to the cross-correlation function as long as the autocorrelation function of the stimulus signal is short enough. The technical solutions for implementing a wideband correlator are given in [61]. For the purpose of this work, the cross-correlation of the received and stimulus signal was done digitally on a separate computing unit.

Dependent on the size of the object of interest, multiple echoes per object

can be detected. Time variant multipath components can be separated easily due to the high spatial resolution as a function of time delay. If the transmitting and receiving antennas are stationary throughout the measurement time, changes in the measured impulse responses indicate presence of moving objects which in turn correspond to moving people. Depending on the application, specific data processing techniques can be used to extract the desired information from these responses. For people localisation, the primary interest is to extract information from the IRF regarding the position of a person. Various position related signal parameters can be estimated from the measured received signal depending on accuracy requirements and transceiver design constraints, such as the received signal strength (RSS), Doppler, TOA, time difference of arrival (TDOA) and angle of arrival (AOA).

Time based parameters provide more reliable and accurate target localisation compared to other parameters as shown in [32, 65] due to the high time resolution of UWB signals. Time-based parameters can be estimated if there is some sort of cooperation and synchronisation between the transmitter and receivers of the system. TOA measurements provide the time it takes for a signal to travel from a transmitter to a receiver, possibly after backscattering from a person or other target of interest. To acquire the TOA, synchronisation and no clock bias is needed between the transmitter and the receiver of the system. TDOA is defined as the difference between the arrival times of the signal to two synchronised receivers. Only synchronisation between the receivers is required for acquiring the TDOA. The TOA, or correspondingly the range between the target and the sensor, is one of the most useful parameters for target localisation when using UWB systems.

It is often suggested that using Doppler information in addition to time-based parameters is advantageous. However in UWB systems the Doppler effect can not be threated as a simple frequency shift, instead the signal is compressed or stretched. This makes Doppler processing complex. Additionally, the systems used in this thesis uses subsampling and averaging

which further complicates Doppler processing.

2.3 EM wave propagation in localization scenarios

If there is only a direct line-of-sight (LOS) between the transmitter and receiver, the received power at frequency f and distance $d_{Tx,Rx}$ between the transmitter and receiver, can be expressed as

$$P_{Rx}(d_{Tx,Rx}, f) = P_{Tx}G_{Tx}(f)\eta_{Tx}(f)G_{Rx}(f)\eta_{Rx}(f)\left(\frac{c_0}{4\pi f d_{Tx,Rx}}\right)^2 \quad (2.3)$$

where P_{Tx} is the transmit power, G_{Tx} and G_{Rx} are the transmit and receive antenna gains, η_{Tx} and η_{Rx} are the efficiencies for the transmit and receive antennas and c_0 is the signal propagation velocity. The received power at frequency f when the signal is reflected by an object at distance d_{Tx} from the transmitter and distance d_{Rx} from the receiver can be expressed as

$$P_{Rx}(d_{Tx}, d_{Rx}, f) = P_{Tx}G_{Tx}(f)\eta_{Tx}(f)G_{Rx}(f)\eta_{Rx}(f)\frac{{c_0}^2\sigma_{RCS}(f)}{(4\pi)^3 f^2 {d_{Tx}}^2 {d_{Rx}}^2} \quad (2.4)$$

where σ_{RCS} is the radar cross-section (RCS) of the reflecting object.

Considering a transmitted signal with flat power spectral density over $[f_L, f_U]$ with $f_U = f_L + \beta$, the received power scattered by the target is

$$P_{Rx} = \frac{P_{Tx}G_{Tx}(f)\eta_{Tx}(f)G_{Rx}(f)\eta_{Rx}(f){c_0}^2\sigma_{RCS}(f)}{(4\pi)^3 {d_{Tx}}^2 {d_{Rx}}^2}\left(\frac{1}{f_L} - \frac{1}{f_U}\right) \quad (2.5)$$

The antenna gains and efficiencies vary considerably with frequency, influencing the received power. In narrowband systems the antenna gains and the object reflection and scattering properties are considered to be frequency independent due to the small frequency band of interest. This

assumption can not be applied to UWB systems. In UWB systems the frequency dependence of material and object properties as well as the frequency dependence of transmit and receive antennas are significant. The antenna efficiency is influenced by the impedance bandwidth of the antenna, specifying a frequency band over which the signal loss is not significant. It is challenging to limit this signal loss to low and fixed levels over a wide frequency band when designing a UWB antenna.

The RCS of an object is a complex function of wave frequency, polarization, aspect angle, and shape of the object. In radar literature, the RCS is considered to be a random variable [66–69]. The RCS generally increases with frequency [70, 71] and can cause significant changes in the received power over the range of frequencies. Its mean power can be estimated using Eq. (2.4). In UWB localization, a person is a complex extended target and its RCS has large influence on the scattering of the signal and thus target detection.

Transmission through an object is another factor that influences the channel characteristics. The dielectric properties of an object affect both transmitted and received signal, and for most materials the dielectric characteristics vary with frequency, for example the dielectric constant of a brick wall has an almost linear increase from 3.7 to 4.48 in the frequency band from 1.31 GHz to 7.01 GHz whereas the dielectric constant of glass, wood and concrete have much smaller variations in the same band [72]. UWB signals consist of many frequency components, some of which can be well reflected from the objects and others can propagate through them. This makes UWB signals useful in many scenarios such as in through-wall applications.

In addition to reflection and transmission, diffraction of an edge and scattering on rough surfaces are two other important propagation characteristics that also depend on frequency. In conclusion, objects in the environment have frequency dependent effects on the propagation and this should be considered when using UWB systems.

2.4 UWB sensor node

A UWB sensor node within this work is defined by an autonomous UWB module with one transmitter, two receivers and a local processing unit as shown in Fig. 2.2. In [26] it is referred to as a bistatic bat-type UWB sensor node. The two receiving antennas are positioned on each side of the transmitting antenna forming a bat-type structure. The transmitter and receivers are synchronised and from each of the two IRFs the TOA of possible targets can be obtained.

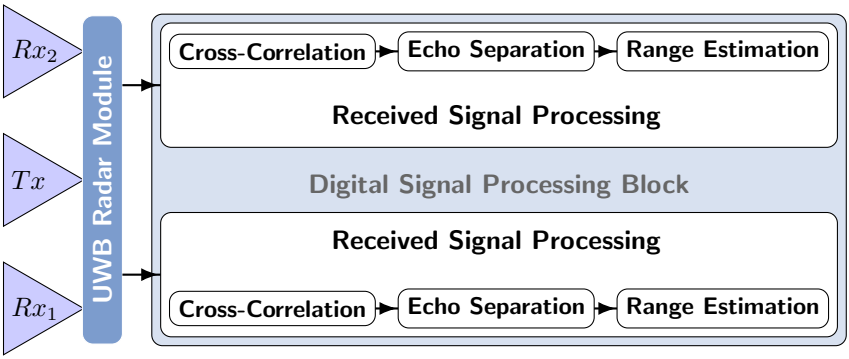


Figure 2.2: UWB sensor node structure

The UWB module consists of the analogue part and the analogue-digital converter. A basic concept is given in Fig. 2.3. More details of the M-sequence radar are given in [40]. The RF clock pushes a shift register which provides the stimulus signal and a binary divider which controls the data acquisition. The shift register generates the M-sequence as defined by its internal structure. Data acquisition is undertaken by sub-sampling due to the periodic nature of the signal, thus reducing the technical requirements of the receiver electronics. For an M-sequence composed of $2^n - 1$ chips, after 2^m signal periods, the binary divider takes care that at least one data sample is taken from every chip of the sequence. According to the sampling theorem, the sampling rate f_c corresponds to usable bandwidth from direct

current (DC) to $f_c/2$. By selecting the RF clock, the desired bandwidth for the system can be matched.

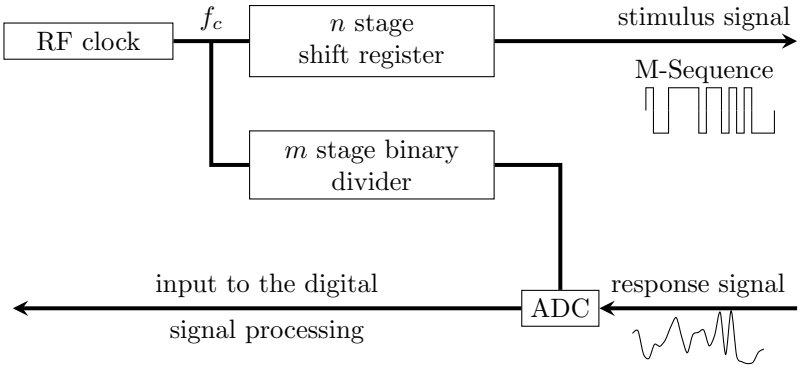


Figure 2.3: Architecture of UWB Radar Module

The sensor signal processing is done digitally. The received signal is correlated with the excitation signal used by the module to define the IRF (cross-correlation). Generally this step is followed by target detection and relevant parameter estimation techniques. For the specific purpose of person localisation the IRF is processed for detecting the scattering of a person (echo separation) and determine this person's range from the sensor node (range estimation). The details for person detection including echo separation and range estimation are given in Section 2.5. Since the sensor node used within this work has two receivers, the range estimates of a person with respect to the two receivers can be used to determine the person location. Details on single sensor person localisation are given in Section 2.6.

A sensor defined in this manner is applicable in many scenarios where the sensor node needs to be used from one side of the scenario of interest such as in through-wall scenarios. A portable sensor that can operate in a "plug and play" mode would be ideal. This means that the sensor should

be quickly ready for use in an unknown scenario and can provide real-time or near real-time information for the presence of people in that scenario.

2.4.1 UWB sensor properties

Each of the two transmitter-receiver pairs of the UWB sensor node can be considered as bistatic radar. The signal received by the receiver contains pulses directly from the transmitter and reflections from targets and obstacles. Two targets can be resolved by the receiver if they are separated in time with a sufficient temporal separation which is defined by the receiver specifications. This theoretical minimum resolvable delay for targets with $\delta(t)$ response is defined as follows:

$$\gamma_\tau \approx \frac{T_p}{2} \approx \frac{1}{2\beta} \tag{2.6}$$

for effective bandwidth β and pulse width T_p. The minimum distance at which a target can be detected is known as the minimum ellipse [73, 74]. A target inside this minimum ellipse can not be resolved by the receiver since its pulse is too close to the direct path from the transmitter to the receiver. The minimum detectable distance is then defined as

$$D_{min} = \sqrt{(x_{Tx} - x_{Rx})^2 + (y_{Tx} - y_{Rx})^2} + \gamma_\tau c_0. \tag{2.7}$$

Sensor coverage is also dependent on the minimum SNR required at the receiver for reliable target detection. The SNR of a target can be calculated as follows [73]

$$SNR = \frac{N_s P_{Rx}}{N_0 PRF} \tag{2.8}$$

with N_s being the number of pulses collected for TOA estimation (also known as scan period), PRF is the pulse repetition frequency and N_0 is the one-sided noise power spectral density. P_{Rx} is the received power after scattering of the target as defined in Eq. (2.5). The limit on received

power due to target reflection is inversely proportional to the product of the distances between the radar antennas and the target [61, 75] as can be seen from Eq. (2.5).

For reliable target detection a minimum SNR, SNR_{th}, is required at the receiver. The maximum distance product as a limit on the bistatic coverage area can be derived using Eq. (2.8) and (2.5) as

$$(d_{Tx,max}d_{Rx,max})^* = \sqrt{\frac{P_{Tx}G_{Tx}(f)\eta_{Tx}(f)G_{Rx}(f)\eta_{Rx}(f)\sigma_{RCS}(f){c_0}^2}{(4\pi)^3 P_{Rx,th}}\left(\frac{1}{f_L}-\frac{1}{f_U}\right)} \tag{2.9}$$

The points satisfying the above equation form a curve called Cassini oval whose shape depends on the distance product [73].

The coverage area of a bistatic radar is thus defined as the area between the minimum ellipse and the maximum Cassini oval as demonstrated in Fig. 2.4, i.e. a target with distance d_{Tx} from the transmitter and distance d_{Rx} from the receiver can be detected by a sensor if

$$\begin{aligned} &d_{Tx} + d_{Rx} > D_{min} \quad \text{and} \\ &{d_{Tx}}^2 {d_{Rx}}^2 < (d_{Tx,max}d_{Rx,max})^* \end{aligned} \tag{2.10}$$

The target detection probability is defined as the probability that a target is inside the coverage area. If directive antennas are used their directionality and coverage area should also be considered for defining the system coverage area. The detection probability can be calculated as the ratio between the coverage area belonging to the surveillance area and the surveillance area itself as defined in [74].

The input of a detection circuit is pure noise in the absence of a target. When a target is present the voltage rises due to the reflection. The time

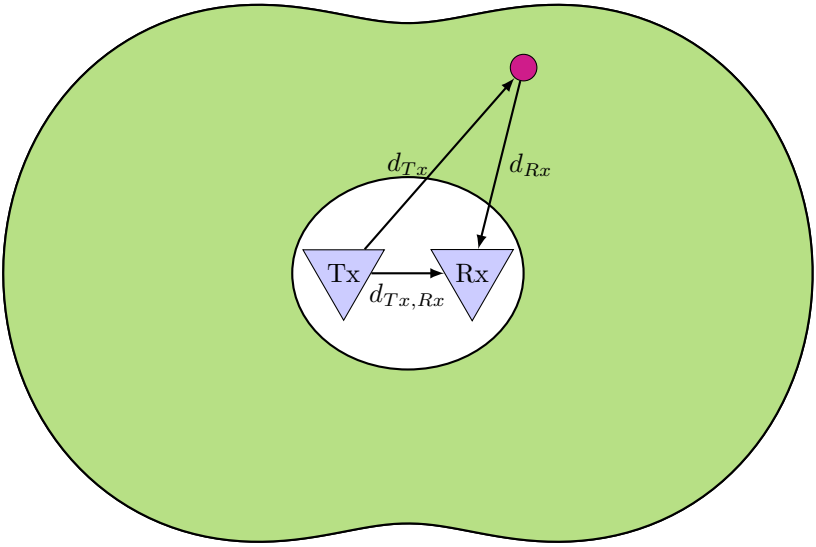


Figure 2.4: Coverage area of an omnidirectional bistatic radar defined as the area between the minimum ellipse and the maximum Cassini oval

position of the voltage above the noise threshold represents the round trip time of the target and its height corresponds to the target reflectivity. Since received signals are often perturbed a signal above a threshold might not correspond to a target and when a target is present the signal might not be above the threshold leading to false alarms and missed targets. The probability of false alarms can be reduced by increasing the threshold level however then the probability of not detecting weak targets increases.

2.5 Person detection principle

Detection of people can be difficult since reflections of a complex target such as a person are much weaker than reflections from metal objects or flat surfaces. Person detection is crucial for person localisation. The received

signal of a UWB sensor contains reflections and scattering from any objects in the environment as well as antenna coupling. All components are also overlaid within one measured IRF. In a typical scenario, the backscattered signal from the people is one of the weakest components in the IRF and the system is subject to many perturbations. The source of these perturbations can come from electronic noise, jamming, stationary and non-stationary clutter, etc.

The reflections of a person are very weak compared to the stationary clutter perturbations. In order to detect the echoes related to reflections of a person, a moving person in a static environment is assumed. Echoes from the person are then time-variant and echoes of other static objects in the environment and antenna coupling are stationary. Separation of the time-variant and static signals is much easier when the person to be detected is moving and not sitting or laying. When the person is stationary, s/he should be detected based on the respiratory and heartbeat activity which is more challenging especially when the person is behind an obstacle. A framework for respiration and heart rate estimation of a single person is presented in [76] and extended for multiple people in [77]. Respiration and heart rate detection is also considered in [77–83]. Experimental demonstration for different antenna polarisation and different body positions is presented in [84] where person breathing cross section is presented.

Knowing the motion profile of the targets (such as person walking speed) it is possible to separate the moving target echo from the stationary clutter (antenna crosstalk and stationary reflections) by a method known as background subtraction as has been demonstrated in Section 2.5.1.

Non-stationary clutter arises from reflections of objects with the same range profile as the desired target or from distant objects, in which case the clutter may be gated out if the unambiguous range of the radar is sufficiently large. Non-stationary clutter is one of the most difficult perturbation sources to be counteracted. The different motion profiles may be used to distinguish between a person and non-stationary clutter. Non-stationary clutter is not

considered at this stage, however [61] suggests that non-stationary clutter often contains strong signal components at frequencies below the breathing rate which can be used to cut the clutter components out.

Once the echoes of the targets are separated from the stationary background reflections, parameters can be estimated that help localise the targets. The restrictions and accuracy of possible parameters which can be estimated are discussed in Appendix II. In this thesis time-based parameters are used. The main sources of error in time-based parameter estimation techniques come from the clock synchronization, the signal acquisition, multipath interference, sampling rate, etc. Precise timing reference is needed between sensor nodes. The different error sources and methods for minimization of their influence in active localization in multipath and NLOS environment are discussed in [5]. Conventionally, TOA estimation is performed by matched filtering or correlation.

To be able to estimate the TOA the most important thing is to determine the start time. In M-sequence radars, a sequence of bits are transmitted continuously in a cyclic manner and thus the start time is not directly known. Shifting to 'time zero' is important for compensation for the length of the RF cables and device internal delays. The reference for this shift can be derived either from additional calibration measurement or from the time position of the pulse related to the direct wave between the antennas. In the second case the TOA parameter of the target is effectively a TDOA parameter where the time difference of the target reflection and direct path is used. For this to work, a LOS direct path is needed between the transmitter and receiver, which in the scenarios and sensor constellation used in this work is not a problem due to the proximity of the two. Since the antenna distance is known, the start time can be easily calculated. The time shifting for determining 'time zero' is done after determining the IRF by cross-correlating the received signal with the stimulus signal of the device used. Within this work the 'time zero' estimation method is used as it is assumed that in unknown and possibly dynamic scenarios a calibration measurement might not be possible.

2.5.1 Moving target echo separation

To detect the weak echoes of the moving persons, the time-invariant strong background reflections should be removed. The first step is to estimate the time-invariant background. Different methods for background subtraction exist [85, 86]. If a measurement of the static background is available, subtracting the background template is a reliable method for target echo separation. In many application scenarios the background cannot be measured ahead of time and thus needs to be estimated. Most common background estimation methods are based on mean, median or mode [87], exponential averaging [88] and low pass filters and predictors [86]. Methods based on mean, median or mode are applicable for offline processing since they require access to all measured IRFs. Exponential averaging can be applied in more realistic scenarios since the background is iteratively computed from the previous background estimation. It is a well suited method for background estimation since it is simple and it is controlled by only one parameter. Background subtraction algorithms that rely on prediction filters can provide more precise adaptation to the specifics of the environment; however they are computationally expensive due to matrix inversion for finding the predictor coefficients.

Due to its simplicity, good performance, high robustness and low computational complexity, exponential averaging is one of the most popular methods for background subtraction. The background estimate as seen by receiver j of sensor s at time t, $u^b_{s,j}(t)$, is computed using the previous background estimate $u^b_{s,j}(t-1)$ and the newly received IRF $u_{s,j}(t)$

$$u^b_{s,j}(t) = \alpha u^b_{s,j}(t-1) + (1-\alpha)u_{s,j}(t) \tag{2.11}$$

with α being a constant scalar weighting (or forgetting) factor between 0 and 1. This factor determines whether recent events ($\alpha \to 0$) or long-term trends ($\alpha \to 1$) in the background estimation are emphasised. For person motion $\alpha \to 1$ should be used since it allows for detection of slow motion in the received signal. The background estimate $u^b_{s,j}(t)$ and the IRF $u_{s,j}(t)$

are both signals in time domain (also referred to as fast time) and the index t is a time index at which the IRF is received. To be complete the IRF should be described using both fast and slow time i.e. the time defining its duration and the time indexes describing the instances in which new IRFs are being received. However for the purpose of the background subtraction the IRF is taken at its entirety with its full duration over the fast time.

The signal of interest containing the person echoes is then

$$u^p_{s,j}(t) = u_{s,j}(t) - u^b_{s,j}(t). \tag{2.12}$$

In Fig. 2.5 the normalised measured IRF, $u_{s,j}(t)$, and the resulting signal after background subtraction, $u^p_{s,j}(t)$, are shown for a given time point t.

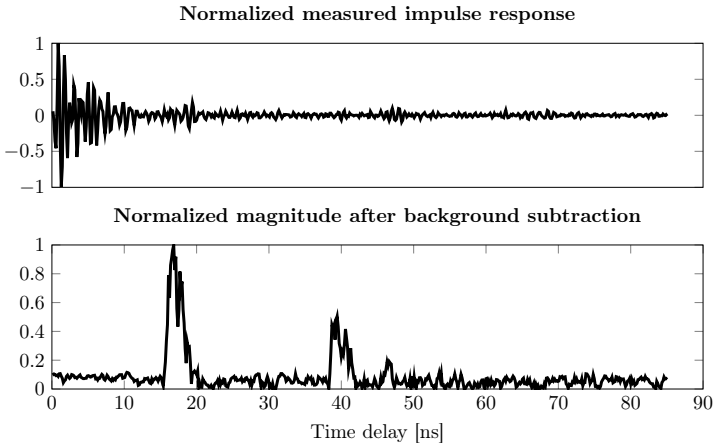


Figure 2.5: Target echo detection - normalised measured impulse response (above) and normalised signal magnitude after background subtraction (below)

In addition to the echoes resulting from the movement of persons in the environment, the IRF contains strong multipath signals from the direct transmitter-receiver feed and the reflections of dominant, flat or metallic

objects. The echoes of the moving people are so weak that they are undetectable in the directly measured signal response. Since the reflections of the background are time invariant and the sensor is stationary, variations in the IRF over time indicate changes in the measured environment corresponding to movement of an object in the environment. After background subtraction the echoes from the moving persons in the scenario are clearly detectable and can be separated from other background reflections and from each other.

The radargrams of people walking in a room before and after background subtraction (over time) are given in Fig. 2.6a and Fig. 2.6b respectively. As can be seen, the reflections induced by both people walking in the scenario can be separated from the static background throughout the scenario duration. One drawback to this method is that it might not be easy to separate closely spaced persons without additional target separation techniques or target detection tracking techniques. The instances where the target further from the sensor is not detectable are discussed later in this thesis and some solutions are proposed (Chapters 3 and 5).

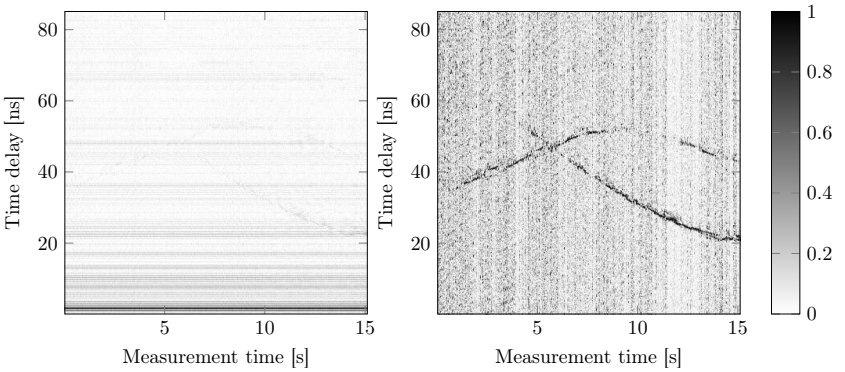


Figure 2.6: Radargram of (a) the raw radar signal and (b) the signal after background subtraction

2.5.2 Range estimation

The signal $u^p_{s,j}(t)$ contains echoes from all time variant reflectors which include the intended targets i.e. people in the coverage area of the sensor. In addition some low peaks in the signal correspond to shadowed background reflections that mimic time variance (see Fig. 2.6b). To extract the range estimates of the targets from the background subtracted signal, various hypothesis or threshold based methods may be used depending on the application requirements. A short overview is given below.

2.5.2.1 Single target ranging techniques

In most indoor scenarios clutter arrives later in the fast time (range) compared to the first target. Thus the estimation of the range of the closest target can be simplified and no clutter models are necessary. The simplest method is peak detection. In this method the first maximum peak of the signal is used to denote a detected target. It has low computational complexity and is suited for completely analogue implementation. However it is only applicable in single target scenarios or when only the closest target to the sensor is of interest. In addition, in multipath channels the first peak of the signal may not be the strongest peak leading to wrong range estimates. Since a person is an extended target, threshold-based approaches are more appropriate [89–91]. Leading edge detection methods with adaptive threshold can also be applied as a relatively cheap and simple method [11].

2.5.2.2 Multiple targets ranging techniques

For multiple target scenarios, more sophisticated methods are needed. Based on the peak detection method, a peak search subtract and readjust method can be applied [29]. Peak detection methods typically require the number of peaks to be estimated in advance and are typically more complex than threshold-based methods. For most multiple target detection methods, the background subtraction data needs to be analysed over a certain period of

time to derive a test statistic and threshold values. Depending on the propagation environment, the choice of the threshold values can be critical for accurately detecting target range. According to the test statistic and threshold, a decision between two hypotheses, H_0 - absence of a target and H_1 - presence of a target, should be taken. The output of the detector is then discrete corresponding to the two hypotheses. Optimum detectors often rely on Bayes, maximum-a-posterior (MAP) or Neyman-Pearson criteria. The most common detectors in UWB systems for person detection are (N, k) detectors [92], inter-period correlation processing (IPCP) [93] and CFAR detectors [94].

CFAR detectors provide good and robust performance for through-wall moving target detection using UWB systems. The threshold level is automatically adapted keeping the given false alarm rate constant. Here we give some more detail on the Gaussian adaptive threshold CFAR detector as used in [95]. It provides the maximum probability of detection for a given false alarm rate based on the Neyman-Pearson optimum criterion. Although the detector is simple and assumes a Gaussian clutter model, it has good and robust performance for many through-wall scenarios. The adaptive threshold is determined using exponential weighted moving average (EWMA) filter. A test statistic X is defined using an EWMA over the unbiased, normalised signal magnitude of $u^p_{s,j}(t)$. The background Y is then estimated using a slower moving EWMA over the signal magnitude and the signal variance σ^2 is defined by using a slow-moving EWMA filter over the signal energy. The adaptive threshold is defined as

$$\theta = \eta\sigma + Y \tag{2.13}$$

where η satisfies

$$P_{FA} = 1 - \int_{-\infty}^{\eta} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\xi^2} d\xi \tag{2.14}$$

for a given false alarm rate P_{FA}. The output of the CFAR detector is then

$$H(X) = \begin{cases} 1 & \text{if } X > \theta \\ 0 & \text{if } X \leq \theta \end{cases} \tag{2.15}$$

The output is a binary sequence which is used to define the time indexes of the signal when a target has been detected (the indexes of the 1's). Each index correspond to the TOA information, $\tau_{s,j}(\mathbf{x}_t)$, of a detected person with unknown state $\mathbf{x}_t$ at time t with respect to the jth receiver of sensor s. The estimated ranges of two moving targets using the CFAR estimator are shown in Fig. 2.7.

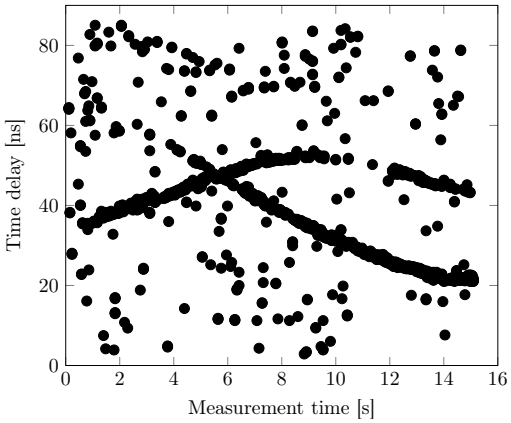


Figure 2.7: Output of CFAR based detector and range estimator

2.5.2.3 Extended targets

Since a person is an extended target when UWB sensors are used for short-range localisation, multiple detections per person are available corresponding to different parts of the person's body that reflected the signal toward the receiver. The extent of a person with respect to the sensor resolution decreases with increasing distance, thus it can happen that people further from the sensors are represented by fewer observations. This has to be considered when modelling the likelihood function for statistically processing the detections. One possibility to deal with the multiple detections is to model or define the extent of the targets of interest. This raises some

difficulties since persons further from the sensors, and in case of multiple targets the targets detected later are represented by much less observations compared to the people closer (or detected first) by the sensor. This can be seen in Fig. 2.9 where the detections of the targets at a given time point are represented in grey. As can be seen the closer target (with TOA of around 20 ns) is represented by more detections compared to the second target (with TOA around 38 ns).

To reduce the multiple detections per target to one, different clustering techniques can be used. Hierarchical clustering [96] with a predefined threshold corresponding to a typical spread of a person in range can be applied since the number of detected persons in the scenario is unknown. However, clustering may hinder detection of multiple targets when they have similar range distance from the sensor. The range detections after CFAR and subsequent hierarchical clustering are shown in Fig. 2.8.

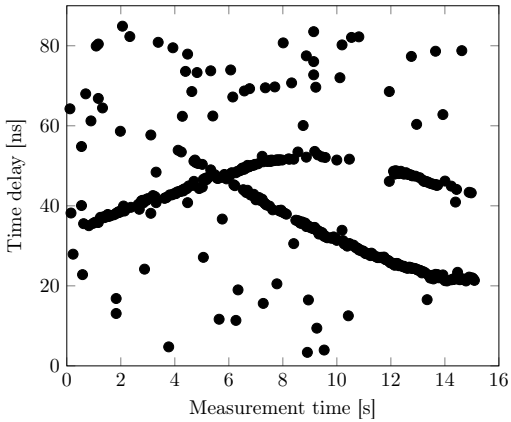


Figure 2.8: Output of CFAR detector after hierarchical clustering

For the obtained results clusters with at least two detections forming the cluster are considered. As can be seen much of the clutter is filtered out.

The performance of the CFAR detector depends on the false alarm rate and the choice of parameters for the EWMA filters. Although these parameters can be adjusted, depending on the position, quality and direction of the sensors, some clutter points will still be classified as targets. These false positives hinder the target localisation and should be removed.

In cases where the leftover clutter and false detections after using the CFAR detector are significantly hindering the localisation of the people, a range tracking algorithm can be applied for removing the clutter and any false detections that do not follow a possible target trajectory. In through-wall scenarios, a wall effect compensation procedure as in [97] can be applied. If the walls are thin with small relative permittivity, the wall effect can be considered negligible.

2.5.3 Tracking of range estimates

To improve the target range estimates and remove clutter, a PHD filter as explained in Section 4.4 can be applied on the CFAR range detections. Different tracking methods may be applied for the same purpose. However we choose a one dimensional Gaussian mixture PHD (GMPHD) filter due to its simplicity and capability to track more than one target. Details on the PHD filter are given in Section 4.4. The target states are defined by the target's range and velocity with respect to the sensor, $\mathbf{x} = [r\, \dot{r}]^T$, whereas the observations used for state update are the CFAR range detections.

As can be seen in the signal sample in Fig. 2.9, the CFAR detector would results in one false target detection around 58 ns in fast time away from the sensor. After applying the simplified GMPHD algorithm, only two Gaussian components remain to represent the states of the detected targets (shown in black), filtering out the false target detection. Applying range tracking over the detections shown in Fig. 2.8, the range estimates from this sensor for the two person scenario can be seen in Fig. 2.10.

When using tracking algorithms for target location estimation, range pre-filtering might not be necessary if the number of clutter and false alarms

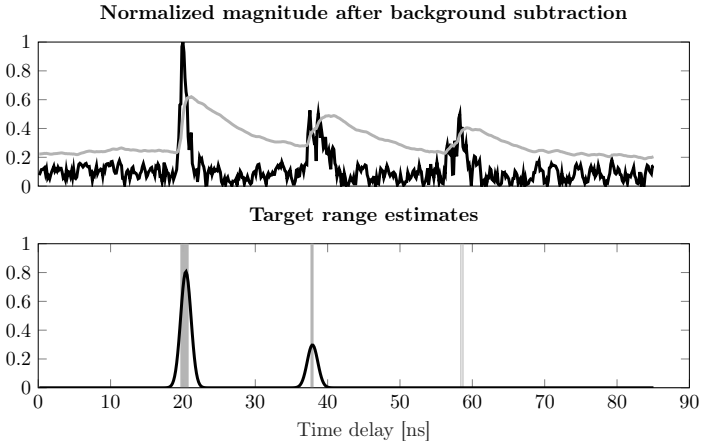


Figure 2.9: Person detection - normalised signal magnitude after background subtraction in black with adaptive threshold in grey (above) and corresponding TOA estimates in grey and GMs from the range tracking in black (below)

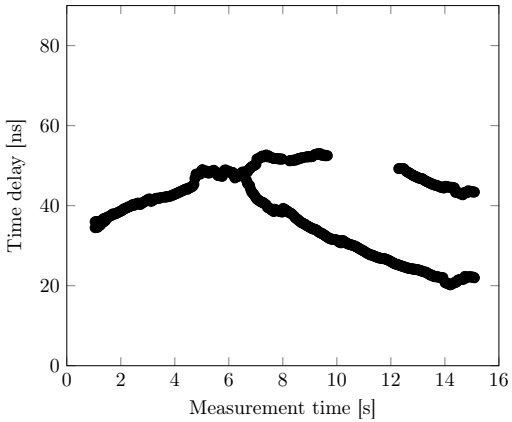


Figure 2.10: Output of CFAR detector after range tracking

is low since the tracking done for location estimation purpose can filter out possible false location estimates in real time. If there are many targets and a lot of clutter the range pre-filtering might be necessary to be able to estimate locations in real time.

2.5.4 Comparison of range estimation methods

In this chapter so far multiple range estimation methods have been discussed. In Fig. 2.11 the relevant range estimation methods are compared on a scenario observed by three different sensors placed around the scenario.

The leading edge detection method as a closest target detection method works well and has little clutter, however is only capable to detect the closest target to the sensor. The second method is the CFAR method with hierarchical clustering as explained above. It performs well, however it may happen that high clutter is present in the signals dependent on the sensor placement. By applying the additional simplified range tracking procedure, most of the clutter is removed from the range estimates.

The advantages of range tracking are especially noticeable for Sensor 3 in Fig. 2.11. For the scenario and movement trajectory considered within this scenario, the sensor placement is suboptimal. Such situations should be expected in unknown scenarios. Due to the suboptimal sensor placement, the sensor can detect the moving targets only in small portions of the scenario run. The rest of the time there are no target-relevant detections.

Which algorithms or methods are applied to process the signals always depends on the scenario requirements. If accuracy in detection of the person is required then most of the methods explained should be applied to assure it. However each applied method uses resources such as time, power consumption, processor resources etc. If the resources are scarce or needed for other applications, some methods such as range tracking, wall compensation, etc may be skipped leading to reduced detection and ranging accuracy. In such cases some false detections are provided to the localisation and tracking framework and should be dealt with there.

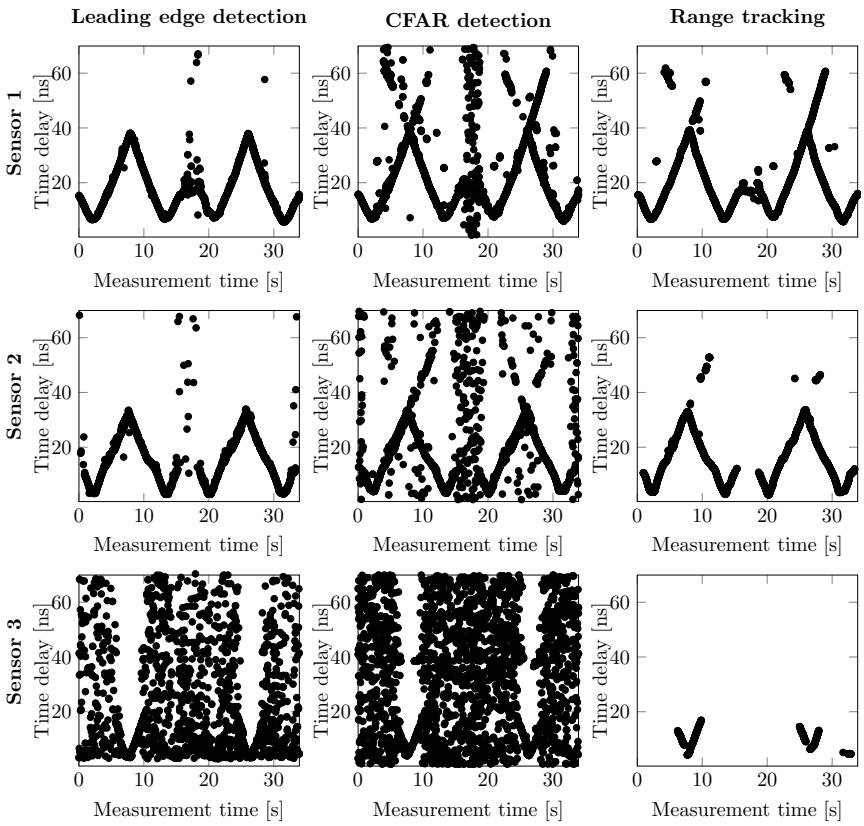


Figure 2.11: Comparison of ranging techniques for multiple sensors looking over the same scenario with two moving persons: (a) the simple leading edge detection, (b) the CFAR detector, (c) the CFAR detector followed by range tracking

2.6 Person localization principle

In single sensor localisation, one UWB sensor node is used to localise and track possibly multiple persons in the area of interest. Ranges are estimated for each detected person in the IRF of each transmitter-receiver pair. The

range information of each detected person by each transmitter-receiver pair can be 'fused' to estimate the number and location of the people in the scenario at a given time.

The number of people detected by each sensor varies over time. Whether a person is detected or not depends on the sensor position, antenna direction and polarisation, as well as location of the people in the scenario. A sensor does not always detect all people in its coverage area. Some are shadowed by other people in the area and cannot be detected. Others may move on the range ellipse of a sensor and appear as not moving to the sensor, resulting in subtraction of their echoes when the static background is subtracted.

In case of multiple persons, having only range information leads to ambiguities in the estimation of the location of the detected people. These ambiguous locations are also known as ghost targets in target tracking literature [98]. In Fig. 2.12 an example of a scenario with two targets and their range estimates is presented. As can be noticed, in addition to the false intersections, four plausible target locations can be extracted. Two of these target locations correspond to the actual target locations and the other two are ghost target locations. In the presence of more targets, clutter and when using more transmitter-receiver pairs, the problem becomes more complex.

Tracking methods can be applied to improve the localisation of the detected persons. In this case, the target location is estimated based on the current range observations of the target and its previous position. In many tracking systems, simple or advanced modifications of the Kalman filter (extended, unscented, etc.) or particle filters are used.

Target localisation and tracking based on UWB radar technology has been previously investigated in [11, 88, 99–105] among others, where [88, 99–101] only consider single target localisation and tracking. Multipath reflections of the person in a known geometry are used to localise a single person behind a corner in [13]. In [102] multiple target tracking using measurement-target data association and a KF bank is investigated. In [11] single target detection per sensor and imaging based data fusion is used for multiple sensor

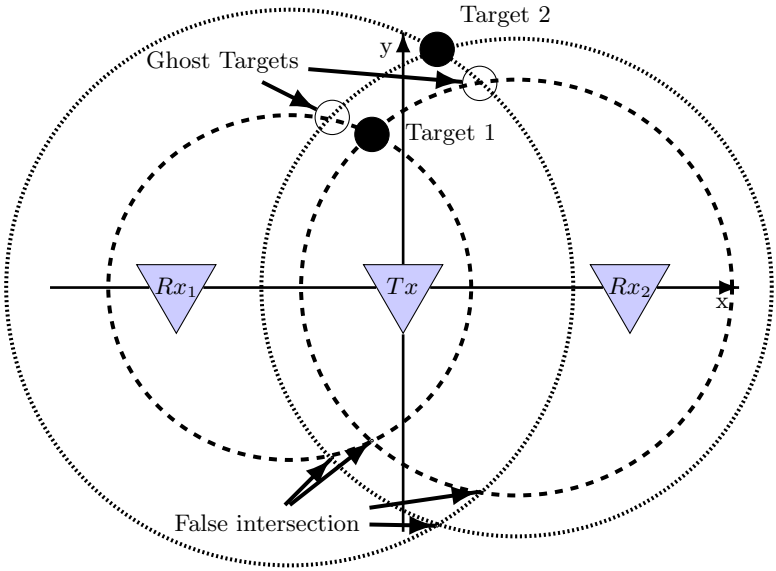


Figure 2.12: Simple scenario with two targets, their range and location estimates with the emergence of two ghost target locations

multiple target localisation. The authors in [103, 106, 107] use multiple hypothesis tracking for both person localisation and characterisation based on parameters estimated from the channel impulse response. The authors in [104] consider people as extended targets and utilise both range and Doppler information for target localisation. The authors in [105] present a TOA based localisation complemented by Cramér-Rao lower bound (CRLB) comparison of TOA and TDOA based localisation.

2.6.1 Range-based localization of a single target

When two receivers are close to each other, it can be safely assumed that a person detected by one receiver would be also detected by the other receiver. The range detection of a person corresponds to an ellipse in two-dimensional

(2D) Cartesian space. This ellipse is defined by the transmitter and receiver positions as foci and the estimated range as the length of the semi-major axis. A person's location can be estimated as the intersection of two ellipses defined by the range estimates from each transmitter-receiver pair of a sensor. The principle of TOA-based localisation using a single sensor is illustrated in Fig. 2.13.

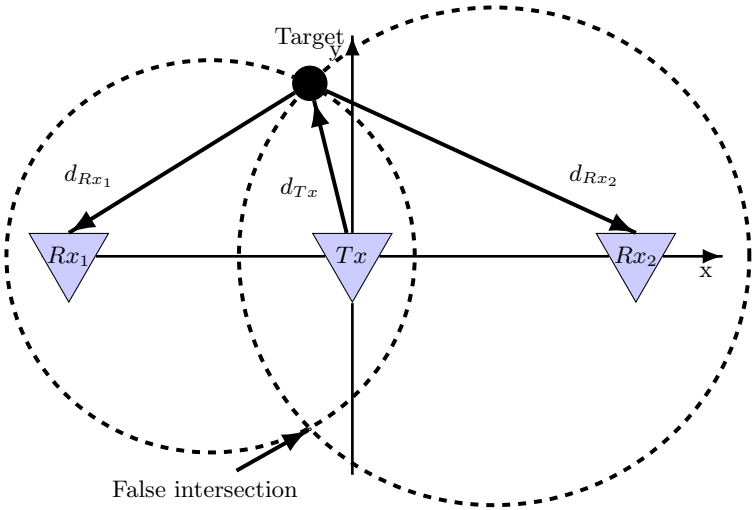


Figure 2.13: ToA-based localisation

The true range of a target $\mathbf{x}_t$ with coordinates (x_t, y_t) at time t detected by sensor s using receiver j is defined as the distance from the transmitter to the target, d_{Tx} plus the distance from the target to receiver, d_{Rxj}, i.e.

$$\begin{aligned} r_{s,j}(\mathbf{x}_t) &= d_{Tx} + d_{Rxj} \\ &= \sqrt{(x_t - x_s)^2 + (y_t - y_s)^2} + \sqrt{(x_t - x_{s,j})^2 + (y_t - y_{s,j})^2} \end{aligned} \tag{2.16}$$

where (x_s, y_s) are the transmitter coordinates of sensor s and $(x_{s,j}, y_{s,j})$ are the coordinates of the j-th receiver of sensor s.

The estimated target range is corrupted by additive noise i.e. random distance estimation error, and can be represented as

$$\tilde{r}_{s,j}(\mathbf{x}_t) = r_{s,j}(\mathbf{x}_t) + \varepsilon_{s,j,t} \tag{2.17}$$

where $\varepsilon_{s,j,t}$ is the random distance estimation error at time t for the transmitter-receiver channel of sensor s and receiver j, i.e. additive noise of the TOA estimation. The target range can also be expressed based on the estimated TOA, $\tau_{s,j}(\mathbf{x}_t)$, of the target using the algorithms presented in Section 2.5:

$$\tilde{r}_{s,j}(\mathbf{x}_t) = \tau_{s,j}(\mathbf{x}_t)c_0. \tag{2.18}$$

The target localisation can then be redefined as estimation of the target coordinates (x_t, y_t) based on the set of non-linear equations (2.17) for all s and j. In the case when single sensor is used ($s = 1$), the solution of two non-linear equations, each corresponding to one of the sensor receivers ($j = 1, 2$) should be found. The same equations can be used for other sensor constellations of multiple input multiple output (MIMO) sensor architectures with multiple transmitters and receivers.

A popular method for solving the set of non-linear equations is based on Taylor Series expansion [108]. The non-linear equations are linearised by expanding the Taylor series around a guess of the solution and only keeping the first order terms. The linear equations are then solved by using the least-squares method and thus estimate the target coordinates. The solution is then used for the next iteration as a point around which the equations are linearised. The iterations continue until a predefined criterion such as minimum error is satisfied.

If a perfect TOA estimate is assumed, i.e. $\varepsilon = 0$, Equations (2.16) and (2.18) can be rewritten as:

$$\tau_{s,j}(\mathbf{x}_t)\, c_0 = \sqrt{(x_t - x_s)^2 + (y_t - y_s)^2} + \sqrt{(x_t - x_{s,j})^2 + (y_t - y_{s,j})^2} \tag{2.19}$$

which is the equation of an ellipse with foci (x_s, y_s) and $(x_{s,j}, y_{s,j})$ and semi-major axis $a_{s,j} = \tau_{s,j}(\mathbf{x}_t)\, c_0/2$ and a semi-minor axis $b_{s,j} = \sqrt{a_{s,j}^2 - e_{s,j}^2}$ where $e_{s,j}$ is the half distance between the transmitter and receiver.

Equation (2.19) can be rewritten as bivariate quadratic equation, i.e.

$$A_j x_t^2 + B_j y_t^2 + C_j x_t y_t + D_j x_t + E_j y_t + F_j = 0 \tag{2.20}$$

where

$$\begin{aligned}
A_j &= (x_{s,j} - x_s)^2 - 4a_{s,j}^2 \\
B_j &= (y_{s,j} - y_s)^2 - 4a_{s,j}^2 \\
C_j &= 2\,(x_{s,j} - x_s)\,(y_{s,j} - y_s) \\
D_j &= 4\,(x_{s,j} + x_s)\,a_{s,j}^2 + (x_{s,j} - x_s)\left(x_s^2 - x_{s,j}^2 + y_s^2 - y_{s,j}^2\right) \\
E_j &= 4\,(y_{s,j} + y_s)\,a_{s,j}^2 + (y_{s,j} - y_s)\left(x_s^2 - x_{s,j}^2 + y_s^2 - y_{s,j}^2\right) \\
F_j &= 4a_{s,j}^4 - 2a_{s,j}^2\left(x_{s,j}^2 + x_s^2 + y_{s,j}^2 + y_s^2\right) + \frac{\left(x_s^2 - x_{s,j}^2 + y_s^2 - y_{s,j}^2\right)^2}{4}
\end{aligned}$$

Since such an ellipse is defined for all transmitter-receiver pairs (s, j) where the target is detected, the target location can be estimated as the intersection point of those ellipses. For one sensor node detecting one target two such equations are defined. The closed form solution can then be derived by solving the 4th-degree polynomial using standard linear elimination of variables (for the case of two ellipse intersections) by applying the Bézout's Theorem.

The number of intersections of two ellipses can be zero if the ellipses never touch, infinity if the ellipses overlap, and one, two, three or four if the ellipses touch or cross each other. Since the two ellipses share one focus (the transmitting node) and the major axis, there are at most two solutions as ellipse intersection points (the problem reduces from fourth to second degree). In the presence of a target two intersection points are expected. Each should have the same likelihood of being the target location, however

the sensor node is typically positioned on one of the edges of the area of interest and directive antennas are used to observe the front half-space of the sensor. Using this knowledge of the sensor-scenario geometry, in most cases one of the intersection points can be directly discarded as undesirable or a low likelihood can be assigned to it since it does not lie within the area of interest.. If a target is present in the scenario and the ellipses do not intersect, it can be assumed that either the target range has not been estimated properly or the target is not detected by one of the receivers of the sensor node and a detected clutter or false alarm point is being used as target information. Same applies to the case of only one intersection point.

A problem with the ellipse intersection method for target localisation arises when the target is far from the sensor. Since the receivers are relatively close to each other, the ellipses are mostly overlapping leading to a flat intersection corresponding to a large localisation error.

2.6.2 Range-based localisation of multiple targets

In the presence of only one target, in theory, there is only one range estimate per receiver, leading to only one possible ellipse intersection for the target location. However, in the presence of multiple targets, multiple combinations of range estimates are possible giving rise to ghost target locations as shown in Fig. 2.12. The calculation of the ellipse intersection points is computationally expensive and time consuming, thus it might be advantageous to first associate range estimates from both IRFs corresponding to the same target.

In the scenarios considered in this thesis the static sensor nodes have both receivers very close to the transmitter and thus a similar range estimate value is expected for a target by both receivers. Based on this assumption an intersection threshold is defined for associating the correct range estimates from the two receivers for the correct target. This intersection threshold is based on the maximum range difference expected of a target detected by both receivers of the sensor node. Since the size of the area of interest

A is known (or can be approximated based on the sensor locations) the intersection threshold, T_s, for each sensor s can be calculated in advance for given transmitter and receiver positions as

$$T_s = \max_{\mathbf{x}_t \in A} |\tilde{r}_{s,1}(\mathbf{x}_t) - \tilde{r}_{s,2}(\mathbf{x}_t)|. \tag{2.21}$$

Let ζ and χ denote two objects in the target space. If the range estimate of object ζ with respect to the first receiver, $\tilde{r}_{s,1}(\zeta)$, and the range estimate of object χ with respect to the second receiver, $\tilde{r}_{s,2}(\chi)$, satisfy

$$|\tilde{r}_{s,1}(\zeta) - \tilde{r}_{s,2}(\chi)| \leq T_s, \tag{2.22}$$

it can be concluded that the objects ζ and χ represent the same object. If multiple range estimates satisfy (2.22), the estimate resulting in the smallest absolute difference is chosen. The target location is estimated as the intersection point of the ellipses defined by the range estimates corresponding to the same target as explained in Section 2.6.1. The above association does not only reduce the computational load, but it also helps in reducing ghost location estimates. Ghost location estimates may result due to intersection of ellipses defined by ranges belonging to different targets, or by choosing multiple intersections of one ellipse with the other ellipses.

2.6.3 Measurement-based method verification

The method described above for localisation of multiple targets using a single UWB sensor node has been verified using measurements from an indoor scenario with two moving people.

Fig. 2.15 shows the estimated locations of two targets moving in a scenario surrounded by four sensors. The different colours of the location estimates correspond to the sensors whose observations were used to determine the location. Fig. 2.14 depicts the estimated positions of the two targets by Sensor 3. The area surrounded by the grey lines denotes the area there the

targets were moving. In Fig. 2.14 the arrows denote the starting position of the two targets and their movement direction.

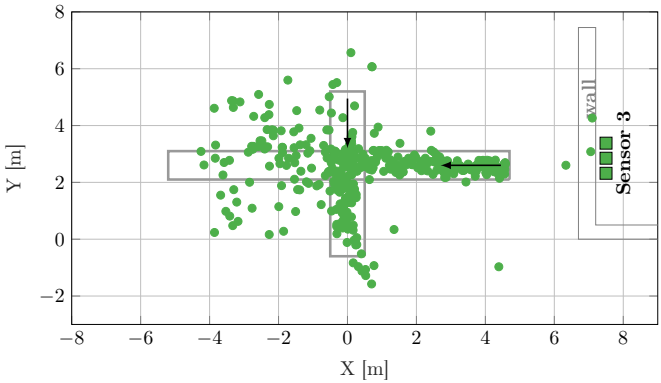


Figure 2.14: Locations of two targets as estimated by Sensor 3. The grey zone depicts area where people were moving.

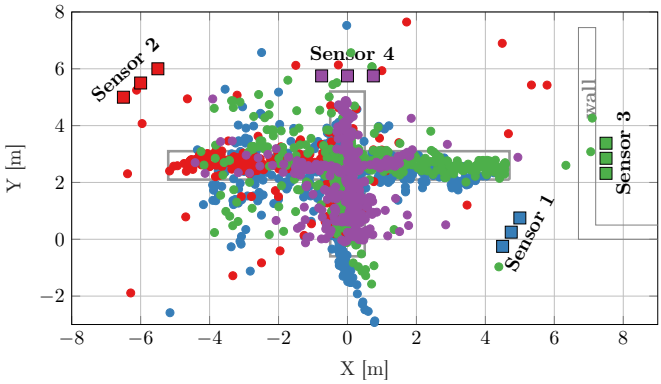


Figure 2.15: Locations of two targets as estimated by four separate sensors. The colour of the location estimates corresponds to the colour of the respective sensor. The grey zone depicts area where people were moving.

As can be seen in these figures, each sensor has different performance in locating both targets over the scenario. This performance depends on how the person is positioned and moving with respect to the sensor as well as the distance of the target to the sensor and whether there is an obstacle (wall) separating the targets and the sensor. As can be seen from Fig. 2.14, the locations of targets closer to the sensor have better location estimates compared to targets further from the sensor due to the flatness of the ellipse intersection with growing distance (leading to inaccurate location estimates).

In Fig. 2.15 the estimated locations by all sensors are shown. As can be noticed, each sensor on its own can not localise both targets continuously and reliably. There are many outliers when the targets are further from the respective sensor. However, all sensors together can cover the scenario and the movement of both targets. Thus as next step a reliable fusion of the location estimates is needed for more reliable target location estimation.

2.7 Concluding Remarks

Indoor localization of persons that do not cooperate with the system network, such as potential intruders, can be complimented by the use of UWB systems. The large frequency spread makes UWB systems applicable for through-wall detection and localization by using the lower frequencies of the spectrum. In this chapter person detection and localisation using UWB sensor system is presented. The principle of UWB sensing and different EM aspects that need to be considered are discussed.

For detection of people in the IRF a combination of exponential averaging for echo separation and CFAR followed by hierarchical clustering is used. The selected methods allow for real-time processing of the received signals and providing accurate range estimates of multiple targets present in the environment. Range tracking procedure is also introduced which can be used when a lot of clutter is present in the detections. However this will

cost some computational resources and some delay in the processing. In well defined and modelled environments, range tracking can be skipped.

Due to the sensor structure with a transmitter and two receivers, a detected person can be localised by finding the intersection of the ellipses defined with major axis equal to the bistatic range of the target with respect to the two receivers. This method is straight forward when each channel provides only one detection. However when reflections of multiple targets are present in the channel, the task is not straight forward any more. The detections from each channel need to be associated with a target so that the correct ellipse intersections can be chosen as target location estimates. If the association is not done correctly false locations or ghost locations may be chosen as target location estimates. In this chapter a method for detection-to-target association is presented and verified on data measured by four different UWB sensors observing an environment with two moving people. The sensor geometry has influence on the accuracy of the location estimates. In the following chapter the concept for distributed UWB sensor network is introduced so that the data from all observing sensors can be used for improving the accuracy of the location estimates.

CHAPTER 3

MULTIPLE SENSOR DATA FUSION FOR LOCALISATION

By using UWB sensors objects in the area of interest can be detected with substantial accuracy, as shown in [109] and above in Section 2.5. A single sensor node is often not well suited for detection and localisation of multiple targets as seen in Fig. 2.14. Targets close to the receiver and/or transmitter of a sensor node shadow over other targets in the environment, making them undetectable or only partially detectable to the sensor.

One possibility to detect those shadowed targets is to use multiple sensor nodes and place them around the scenario. In such case at least some of the sensors can detect some of the targets in the scenario. By fusing their detections appropriately all targets may be localised. This chapter describes the concept for a distributed UWB sensor network taking into consideration the issues that need to be addressed when designing such a network for real time operation. Possible sensor operation principles, system synchronisation and sensor cooperation issues are discussed, followed by description of the sensor network architecture with sensor node and fusion center frameworks. After discussing some real time operation and implementation challenges, the multiple sensor fusion concept is presented. Different challenges present in multiple sensor fusion scenarios are discussed and a multiple sensor likelihood function is derived. The likelihood function is then visualised

and compared based on the input provided for a two person scenario.

3.1 Introduction

In a scenario with reflective walls and multiple objects, the *background subtraction* method may fail in separating the moving from static objects. One of the reasons is the unavoidable interaction of the multiple reflections and especially the influence of shadowing [10] either by static objects over the targets, or dynamically by the targets themselves. As a result, the SNR is low, leading to multiple false alarms that influence the detection and localization accuracy. Since the estimated parameters are also noisy, and in the UWB sensor structure used here the two receivers are close to each other, the estimated position is distorted. Additionally, sensors have limited field of view and limited accuracy in ranging depending on the target distance from the sensor. Due to the use of directional antennas, blind-spots arise in the scenario where a sensor can not detect any present target.

Distributing multiple UWB sensors around the scenario decreases the geometric dilution of precision and increases the coverage area of the system. Different sensor distribution schemes for UWB systems have been discussed in literature [26, 109, 110]. The distributed UWB sensor network scheme used here is based on multiple autonomous bistatic bat-type UWB sensor nodes (described in Section 2.4) placed at any convenient[1] locations around an area of interest. Fig. 3.1 demonstrates the idea of the sensor network scheme. Multiple standalone UWB nodes are distributed around an unknown area of interest and each is connected to a common fusion center or control system which provides the target detections and locations to the system user.

A scheme like this is useful since no operators are needed around each sensor node at all times. The sensors need only be placed at the desired location and the full control of the sensors can be done from the fusion

[1] In this case convenient is meant as spatially convenient placement and not due to optimal sensor positioning. Optimal sensor positioning is not in the scope of this work.

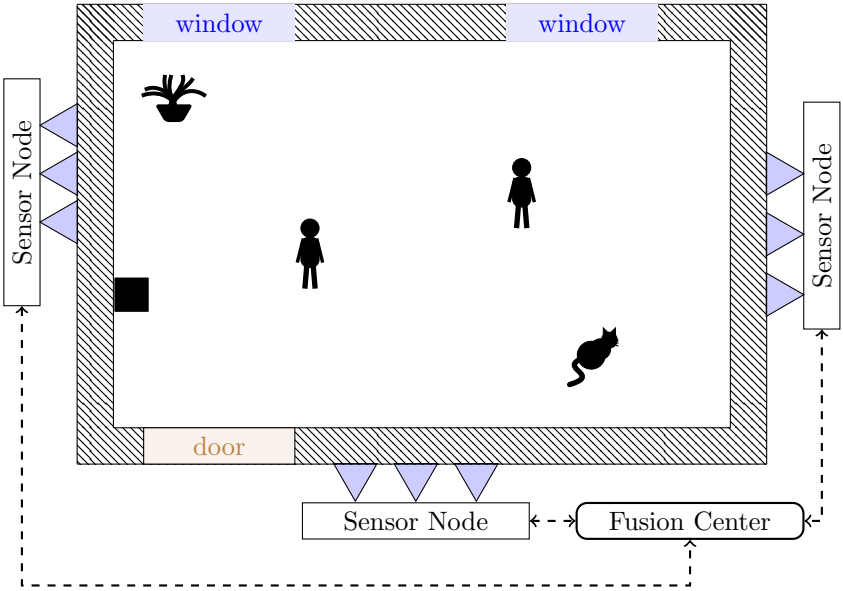


Figure 3.1: Example of possible scenario of interest surrounded by multiple UWB sensor nodes

center by a single operator. The network covers a larger area and many shadowed regions and blind-spots in the area of interest with respect to one sensor are covered by another sensor of the network. Moreover not all sensors are influenced by multipath in the same manner. Targets in the area of interest may be detected by multiple sensors of the network, providing information regarding the target from different angles. This information needs to be properly fused accounting for target extent, ranging errors, bias and systematic errors so that a target detected by two different sensors does not appear as two separate targets close to each other in the final results.

When using distributed sensor networks, it is desired that as little data as possible is transmitted between nodes. Thus sending the full channel

impulse response to the central node compared to few parameters relating to the detected targets might be excessive. Additionally, attention must be paid to the performance of the overall system in real time. Both the detection of the targets by the sensors and the fusion of this information must be provided in real or near real time.

Person detection and localisation using distributed UWB sensor networks have been studied in [26, 73, 74, 111–113]. In [111] the CRLB for object tracking in a 2D environment assuming specular reflections is derived. The authors in [73] study the impact of the sensor network geometry on the target localisation for anti-intruder UWB radar systems. Optimum detection metrics for target detection and tracking in UWB sensor networks are defined in [112]. The authors in [74] analyse UWB sensor networks, taking into consideration the sensor coverage area and probability of target detection.

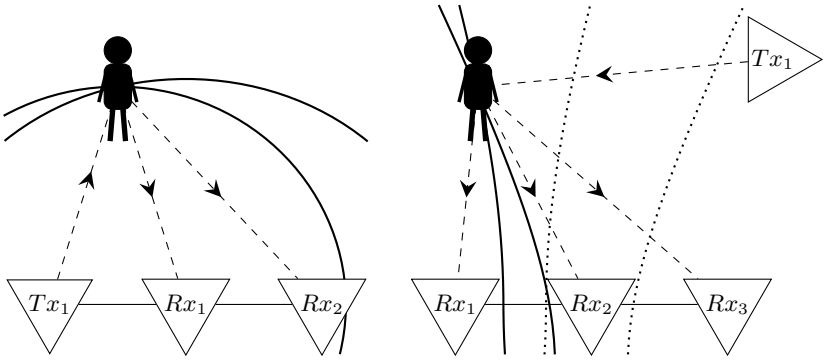
Within this chapter the bat-type UWB sensor is considered as a single sensor node and multiple such nodes are used to design a sensor network around scenarios of interest. The considerations taken when designing a distributed UWB sensor network are described further in this chapter, followed by derivation of the multiple sensor likelihood function for fusion of the data from the contributing sensors.

3.2 Sensor network design and sensor cooperation principles

Using distributed sensor network has many advantages, and can significantly improve the UWB radar system abilities due to the diversity gain offered by the sensor network. However, to exploit this, we need appropriate sensor cooperation and synchronization, as well as appropriate fusion of the information obtained by each of the sensors in the network. For coherent real-time signal processing, the components of the UWB sensor need to be temporally synchronized.

3.2.1 System synchronisation

Synchronisation of the transmitters and receivers in distributed sensor network is necessary for acquiring the target range and location information. To obtain TOA information about a target of interest a transmitter and receiver need to be synchronised. For TOA localisation three synchronised transmitter-receiver pairs are required. This number can be reduced to two if additional information about the observation direction of the sensors is provided. An example is shown in Fig. 3.2a where one transmitter is synchronised with two receivers. By use of directional antennas the observation direction is also known, and can be used to calculate a target location. Depending on the scenario and network constellation, it might be simpler to have only two synchronised receivers which are not synchronised with a transmitter, for example a scenario where the receivers can be placed on the same platform and a transmitter needs to be placed at a distance. In this constellation TDOA information can be used to determine a hyperbola of the range of the target as in [109]. If at least three temporally synchronized receivers are provided, two hyperbolas whose intersection would result in the target location as in Figure 3.2b can be obtained.



(a) TOA based localisation **(b)** TDOA based localisation

Figure 3.2: Target localisation methods

When bat-type sensors are used there is inherent temporal synchronisation between the transmitter and the two receivers of the sensor since a local RF clock is used (the receivers of the platform are also synchronised). The TOA information from each transmitter-receiver pair can be used to determine the target location as in Section 2.6. Since the receivers are synchronised, using the transmitted signal of a transmitter from a separate sensor node, a TDOA parameter may be estimated as well. This requires knowledge of the transmitted sequence by the external receivers for appropriate correlation when determining the IRF. An example of a scenario where one transmitter from a bat-type sensor node is used is shown in Fig. 3.3. Using the receivers of the same platform two TOA parameters, each from the respective receiver, can be estimated as target observations. If another bat-type sensor node is present or additional two synchronised receivers are present in the scenario, a TDOA parameter may be estimated and used in addition to the TOA parameters as additional information which may improve the target location estimation.

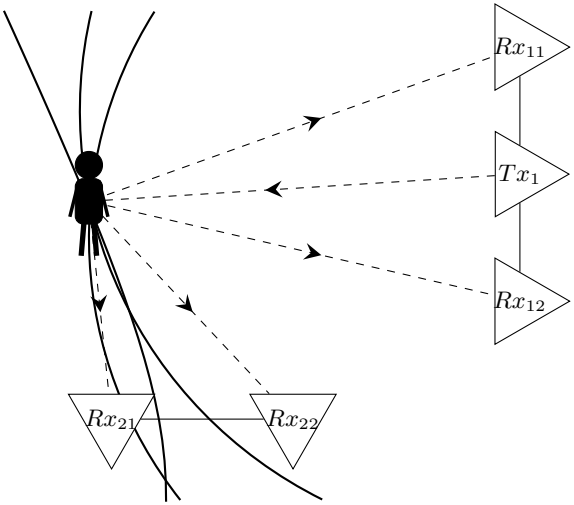


Figure 3.3: Target localisation with TOA and TDOA using UWB sensor nodes

The connection between each sensor node and the fusion centre node is either wired or wireless. To synchronise the system, a common system clock should be shared among all sensor nodes either via coax, optical cables or wirelessly. The wired synchronisation offers more precise and stable measurements, however it requires the use of long cables. The synchronisation between the sensor nodes and the fusion center is needed so that the provided data may be appropriately fused. The synchronisation accuracy here depends on the target speed and does not have to be in pico- or nano-second range as is required for TOA estimation. For person localisation millisecond time accuracy between the sensor nodes and the fusion center is sufficient.

3.2.2 Sensor network operating principles and control

How the sensor nodes cooperate with each other and how can they be controlled is another network design question that needs to be addressed here. In a sensor network defined by using multiple bat-type UWB sensor nodes, two operating principles can be defined: sequential and parallel.

The sequential network operation mode would reduce the interference caused by multiple transmitters being active at the same time. However this increases the measurement time due to switching through the transmitters used in the network. For the sequential mode, synchronisation or a shared system clock between all sensor nodes is required, otherwise the switching between the different sensors may be too slow causing latency in the data delivered to the fusion center and thus resulting in localisation errors. In the sequential operational mode, the master node has a greater role. It can control when a specific sensor node is active and avoid having multiple sensor nodes being active at the same time. A sensor scheduler needs to be implemented with appropriate scheduling algorithm. In this case the active sensor node can push target detections to the fusion center to improve on the location estimation by fusing this data over time and with data from other sensor nodes The target detection can be either range estimates from the two receivers or locally calculated locations. Another possibility is to only

control the transmitters of each sensor node and have all receivers active at all times. The sensor whose transmitter is active can use the signals received by its two receivers to estimate the TOA of possible targets, whereas the other sensor nodes can estimate the TDOA of possible targets. In this case more information can be gathered since in addition to the TOA information from the active sensor node, TDOA information can be gathered by all other sensor nodes whose transmitter is currently not active (see Fig. 3.3). The master node will have to have a dedicated and accurate transmitter scheduler implemented that would also consider possible communication delays.

In a parallel operating principle RF synchronisation between the sensor nodes is not required. By relaxing the synchronisation requirement, each sensor node operates using a local system clock. The transmitters of each sensor node may operate in parallel, and the signals scattered by the environment can only be resolved by using the receivers which are synchronised with that transmitter, given that the transmitted signals of the different transmitters are orthogonal. This means that for each transmitter of a sensor node, only the two receivers of the same sensor node are used. Since all sensors are active at all times, there is not much control from the master node (also known as the fusion center) over the sensor nodes. Basically the master node is only used for gathering the data from the sensors and processing it for estimating target locations. The advantage of the parallel mode is the reduced measurement time for acquisition of the 'reduced' MIMO impulse response function and the lack of synchronisation between the sensor nodes.

Within this work the parallel operational mode was used. Although the sequential operational mode was considered, especially due to the information gain, it was decided that the higher measurement rate of the parallel mode is more important. Additionally, the single sensor processing described in Chapter 2 can be reused and single sensor failure in the network should not cause problems to the network operation. In the conducted experiments the influence of the interference of the other transmitters was

not significant which also meant that even if sequential mode was used, TDOA information could not be obtained (the obtained IRF when using transmitter of a different platform resulted in a noisy signal most likely due to frequency mismatch in the RF clocks). Another practical advantage of the parallel operating mode is that single sensors can easily be disconnected from the network and used autonomously or moved to a more optimal position and connected again.

3.3 Sensor Network Architecture

Although there is no synchronisation between the different sensor nodes, each sensor node cooperates with a central master node which controls the operation and data flow of the sensor nodes. The fusion center node is responsible for sensor discovery and definition of the sensor network based on the number of connected sensors. It also provides a common reference coordinate system based on the knowledge of the sensor locations. The sensors are assumed stationary and with known location, i.e. they are capable of self-localisation. If a sensor is moved to a different area, it is deactivated from the network and reconnected as a new sensor node on its new location.

3.3.1 Sensor node framework

The structure of the UWB sensor node has been described in Section 2.4. As mentioned, the sensor node consists of a stand-alone UWB module with a processing capability, i.e. connected to, for example, a mini-PC which acquires the data from the UWB module and can process them locally before sending processed data to the fusion center.

The framework of a sensor node is shown in Fig. 3.4. The **Sensor Driver** block provides local access to the UWB module for data acquisition from the UWB module and propagating sensor control messages to the sensor itself if specific transmitter or sensor control is desired for the specific scenario. These control commands may also define specific sensor parameters such

as subsampling rate, sensor measurement rate, etc. It can also be used to provide access to a simulated sensor module.

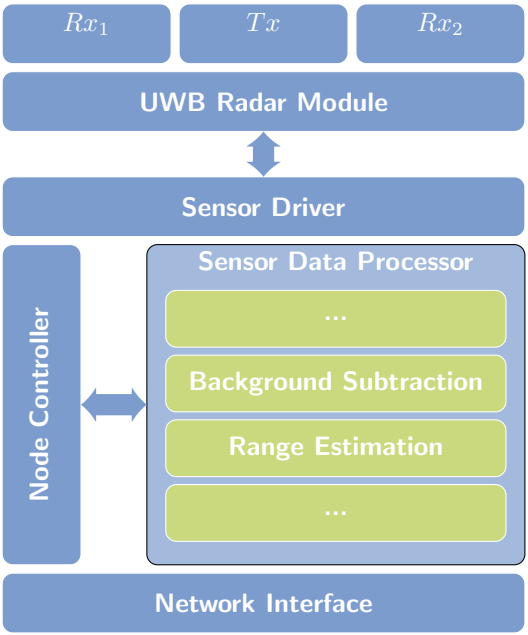


Figure 3.4: Sensor node structure

The **Sensor Data Processor** block processes the received UWB pulses by both receivers to provide range estimates of the detected persons. It consists of the processing steps defined in the green building blocks. An additional processing step for **Target Localisation** can be defined where the ranges estimated by the two receivers are fused to result in locations of the detected targets as explained in Section 2.6. The **Node Controller** block provides the interface for the data structure and transfer from the module to the processor and the processed data to the network interface. It provides data buffering capability from sensor driver to data processor,

and from data processor to the network interface, pushes remote sensor configuration and control messages to the sensor driver and provides remote access to buffered data via the network interface module.

Finally, the **Network Interface** block provides communication between the sensor node and the fusion center node. It has a data interface for transmitting the processed data by the sensor node to the fusion center node, a configuration interface for configuration of various parameters needed for the algorithms used for data processing and a control interface for providing controls for the sensor operation from the fusion center node.

3.3.2 Fusion center framework

The fusion center acts as a control center for all sensor nodes. It manages the operation of each sensor and is responsible for sensor discovery and definition of the sensor network. It has a sensor scheduler which controls the measurement operation sequence of all sensor nodes in the network and defines the operation mode of the sensors. In the parallel operation mode the sensor scheduler is not used, i.e. all sensors operate in parallel. The data received by each sensor is first aligned to account for communication delay and missing data. Then it is passed to a processing unit where it is fused defining location estimates and tracks for each of the detected targets in the scenario. Depending on the design of the network, either locations or range estimates are sent to the fusion center. For fusion of the range estimates of each sensor into location estimates, a common reference coordinate system is defined. This means that the transmitter and receiver locations of each sensor with respect to a common reference node should be known to the fusion center. The framework of the fusion center node is shown in Fig. 3.5.

The **Network Interface** block provides a management interface for managing the operation of each sensor and a push/pull data interface for gathering the data from all sensors. It is equivalent to the network interface block of the sensor nodes. The **Sensor Manager** block implements the

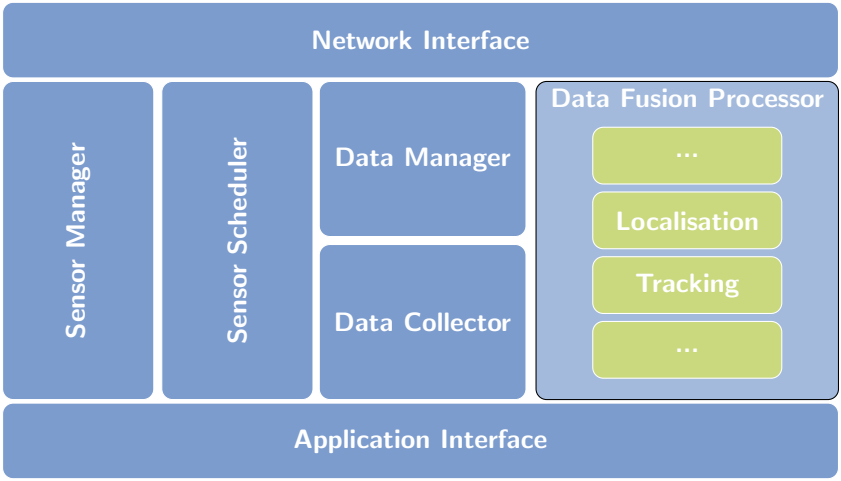


Figure 3.5: Fusion center structure

sensor group management, i.e. sensor discovery. The **Sensor Scheduler** controls the sensor operating sequence and defines the operation mode of the sensors. It uses the control interface of the sensor nodes. The **Data Collector** block acquires and aligns the data received from the sensor nodes and passes it to the processing unit. The data alignment uses the signal acquisition time provided together with the target detections. Although all sensors are not synchronised with a common master clock, they all use distributed clock synchronisation protocol. The **Data Manager** block provides access to current and recent system data using the data structure required by the **Data Fusion Processor**. The data processor consists of different processing steps (represented in green) for fusing the target detections provided by all sensors and thus estimating more accurate target locations. Target tracking is also inherently applied. The target tracks are then provided via the data manager to the **Application Interface** which can then provide them to any application that may use them, for example a visualisation application.

3.3.3 Distributed UWB Person Localisation Demonstrator

Based on this work a demonstrator was built and tested in different environments as well as presented live during the Future Security conference (Bonn, 2012) in its basic form. Later on the experimental demonstrator was upgraded with the newly implemented data processing algorithms.

For the design of a sensor network, multiple challenges are faced. Here it is assumed that the sensor nodes are at known locations with respect to each other. This could be done using global positioning system (GPS) and each sensor node can be self-localised, or active target localisation techniques and time difference of arrival method can be used to localise the sensor nodes [3, 114] (for TDOA based active localisation at least three synchronised receiver antennas are needed to localise a transmitter). If GPS can be used for sensor localisation, it can also be used for synchronising the time between the different sensors. The reason methods other than GPS are investigated for sensor self localisation and time synchronisation, is because this sensor network for person localisation is meant to be used indoors and possibly in areas where the GPS signal might not be reliable or even available. Additionally, the accuracy of GPS localisation might not be sufficient for the small size of the scenarios of interest. Thus local network based time synchronisation is used and active UWB transmitter localisation is considered for sensor localisation.

An experimental system for on-line real-time processing of data-streams by multiple distributed UWB sensors for localisation and tracking of multiple non-cooperative targets was designed and implemented according to the design layout described in the previous sections of this chapter. An example scenario setup is shown in Fig. 3.6. The UWB modules used provide an Ethernet interface with two TCP/IP sockets: a command socket where an application can configure and control the module (e.g. transmitter operation) using binary message protocol, and a data socket where an application can read the received signals of the receivers as stream of 16 kB length. The data socket provides the readings of the sensor's receivers to the sensor data

processor where they are directly processed. It also provides the sensor position coordinates to the fusion center for reference plane definition.

The fusion center provides an interface for recording data from experiments for off-line analysis and an interface for various user applications for visualizing the processed sensor data in real-time. It also provides control and data interfaces similar to the sensor node for access of sensor parameters and current tracking data such as target locations, confidence factors, etc. The sensor nodes does not need to have a physical connection with each other or with the fusion center. Both Wireless and Ethernet based communication between nodes and the fusion center were tested. A decentral service discovery and network-node auto configuration based on Zeroconf [115] is used for sensor discovery. The processing algorithms are implemented in MATLAB, while the rest of the sensor network system is implemented in JAVA. For simplification of further extension and usage of the system, usage of low-level network or operating system interfaces is avoided. Internet communication engine [116] is used for remote interfaces due to its simplicity and portability.

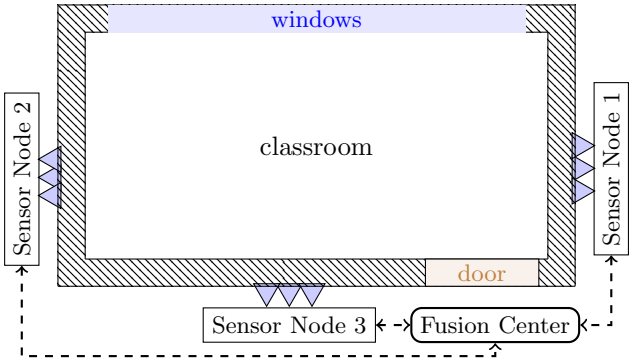


Figure 3.6: Schematic of a measurement scenario setup

The sensor nodes are distributed around the area of interest and as soon as they connect to the network, they are registered on the fusion center side.

In the scenario depicted in Fig. 3.6 three sensor nodes are placed behind each of the walls of a classroom (Fig. 3.7) as can be seen in Fig. 3.8. Each sensor node is driven by a separate 7 GHz RF clock and the measurement rate is set to about 100 IR/s. Using this setup multiple scenarios are measured, using different trajectories and different number of persons as targets, as well as scenarios where the classroom is either full or emptied from its furniture. Some results on the person localisation improvement when using multiple sensors are shown in Appendix III for one and two persons walking in the classroom.

Figure 3.7: Classroom as area of interest

(a) Sensor 1 **(b)** Sensor 2 **(c)** Sensor 3

Figure 3.8: Sensors as placed around the area of interest

3.4 Multiple sensor data fusion

Sensor data fusion describes the combination of data either from multiple sensors or from a single sensor such that the result of it has lower uncertainty compared to when the sensor data is used individually. In multiple sensor fusion the data provided by all sensors observing the desired targets should be fused to accurately represent their actual states. The theoretically optimal method for fusing data from multiple sensors is to fuse the observations from all sensors in a central fusion center [50]. In fusion and tracking approaches, the observations are associated to either an existing track or used to initialise a new track. More on data fusion and tracking approaches is discussed in Chapter 4.

Sensor data fusion consists of three main tasks: data association, target detection and target localisation. The data association task identifies which observations from different sensors belong to the same target and which are false alarms. The target detection determines how many targets are detected and which sensor provides observations for the detection of specific targets and the target localisation uses the provided observations to estimate the locations of the detected targets.

As mentioned in Chapter 2, each sensor is capable of estimating the locations of detected targets. A target is typically detected by more than one sensor as seen in Fig. 2.15. It can be built up upon this method by fusing the locally estimated target locations by the different sensors. Any target detection that can not be used locally for estimating a target location is discarded. Within this work this method is also referred as distributed fusion approach. In a centralised fusion approach all target detections (range estimates) are fused together at the fusion center. Target detections which can not be used locally for target location estimation are not discarded. Instead they are used for target location estimation in the fusion center, where they can be paired with target range estimated from different sensors. The two fusion methods and their differences are also separately explained in Appendix IV.

Association of the observations made by the different sensors to the correct targets is one of the main challenges since the number of possible associations increases the more sensors and targets are present. All targets are not detected by all sensors at every time step. There are many non- and missdetections as the target moves further from a sensor, and if detected the accuracy of this detection is degraded with distance.

3.4.1 Multiple sensor fusion challenges

When using multiple sensors for detection and localisation of target, different challenges can be expected. Some of these challenges arise due to the parameter estimation errors and sensor bias. Fusion of inaccurate and biased parameters estimated in a state space different from the localisation state space can be challenging and thus an appropriate estimation error representation is needed. Within this section some of these challenges are discussed. For simplicity ellipses are used to represent target TOA detections in Cartesian state space instead of likelihood functions or annuli.

In the ideal case with no TOA estimation error, the target location estimated by each sensor coincide resulting in a scenario as seen in Fig. 3.9(a). However, in practical applications there is always a non-zero estimation error. The case when the range estimation error is small is shown in Fig. 3.9(b). It can be observed that the target location estimates with respect to each sensor (blue point intersections) are close to the true target location (black point intersection). In some scenarios, such as through-wall, the range estimation error is large leading to the situation depicted in Fig. 3.9(c) where the target location estimate using sensor 1 is further from the true target location. Finally, it can also happen, in addition to TOA estimation errors, that a target is not detected in one of the IRFs of a sensor as depicted in Fig. 3.9(d), where a target cannot be localised by a sensor since it is not detected in one of the transmitter-receiver channels of the sensor.

These cases have also been considered in [117] where single target localisation by two UWB sensor nodes is considered. A joining intersection of

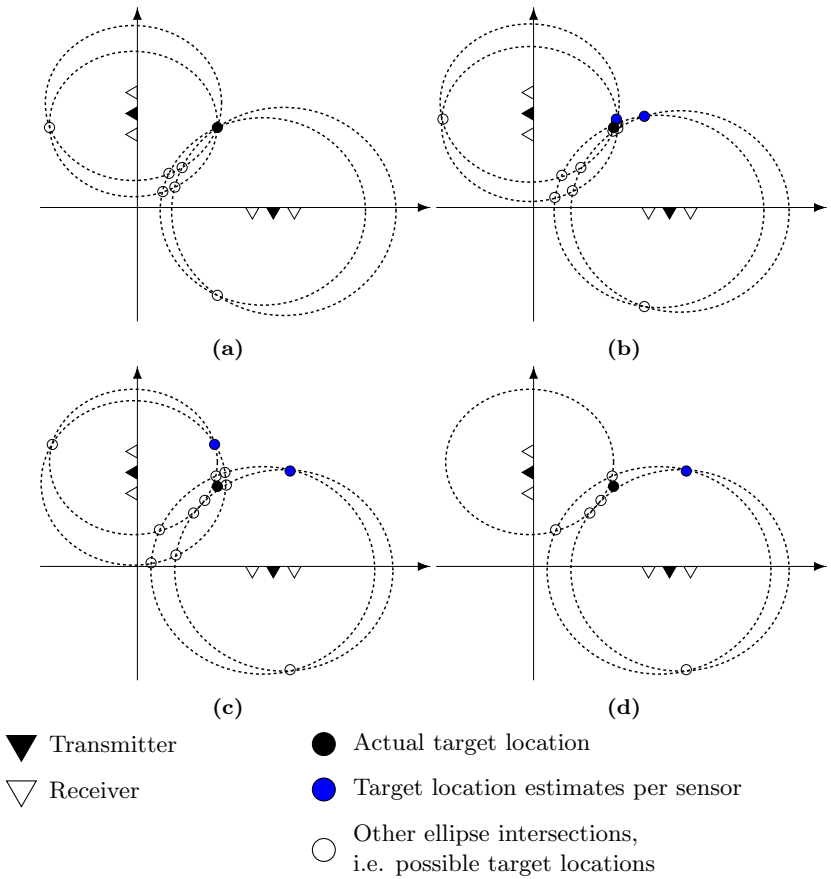


Figure 3.9: Range estimation error influence on location estimates in multiple sensor scenarios: (a) Ideal localisation, no range estimation error; (b) small range estimation error; (c)large range estimation error by one sensor; (d) target missdetection by one receiver of one sensor

ellipses method is proposed, where additional ellipse intersections close to the true target position are utilised to form a cluster of intersections, which

is later used for target positioning. The presented approach is shown to be better than simple averaging of locations estimated by single sensors or Taylor series methods. However, it is also computationally extensive since all possible ellipse intersections should be determined to form the cluster of intersections.

In the presence of multiple targets and high clutter, this method would be impractical mainly due to the many ellipse intersections that need to be determined, the measurement to target association involved as well as the possibility of not being able to correctly separate multiple targets. In multiple target scenarios, it can happen that the cluster of intersections determines a false target location instead of two separate targets, as shown in Fig. 3.10 where one target is detected by one sensor and the other target is detected by the other sensor. If it is not known how many targets are present in the scenario, the cluster of intersections determines one target location instead of the expected two targets.

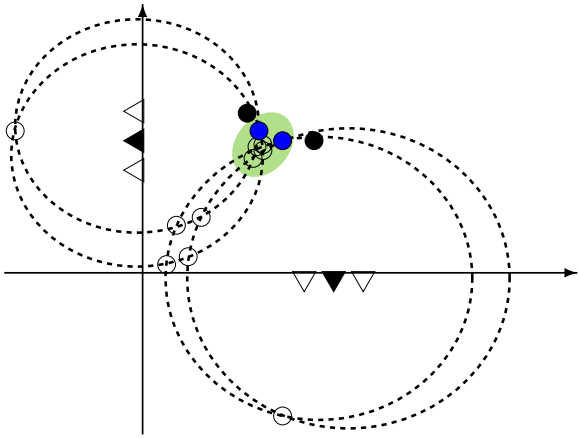


Figure 3.10: Possible location estimation error source in multiple sensor multiple target scenarios: false intersections may be chosen as single target location (green ellipse) instead of the true locations of two targets (black circles)

An additional challenge not described in the above cases is the detection of clutter by some of the transmitter-receiver pairs. This becomes a problem when there is an unknown number of multiple targets and many sensors placed around the scenario for target detection.

In the next section the likelihood function is derived for multiple sensor scenarios, and some instances of it based on the acquired scenario observations are presented.

3.4.2 Multiple sensor likelihood function

A likelihood function characterizes the information obtained from sensors. For the purpose of this work, a joint likelihood function as described in [118, 119] is defined using the range information from each transmitter-receiver pair, the approximate size of the area of interest and the predefined Cartesian coordinate system. The classical likelihood function in the presence of K observations, $\{\mathbf{z}_{t,k}\}_{k=1}^{K}$, from one sensor (sensor pair $[sj]$) for a single target can be defined as

$$p(Z_t{}^{[sj]}|\zeta) = \prod_{k=1}^{K} p\left(\mathbf{z}_{t,k}|\zeta\right) \tag{3.1}$$

where

$$p\left(\mathbf{z}_{t,k}|\zeta\right) = \mathcal{N}\left(\mathbf{z}_{t,k}; h_{s,j}(\zeta), R_t{}^{[sj]}\right) \tag{3.2}$$

represents the functional relationship between the target state $\mathbf{x}_t$ and observation $\mathbf{z}_{t,k}$ as Gaussian probability density function (pdf) with observation noise covariance matrix, $R_t{}^{[sj]}$, at time t. $h_{s,j}\left(\zeta\right)$ denotes the measurement function defining the relationship between the observation and the target state for sensor $[sj]$. Equation 2.16 is used as measurement function since range is the only target observation used in the scenarios considered within this thesis. However the equations presented here can be used for any combination of multiple target observation parameters as typical for Radar observations. The noise variance is typically different for each estimate[2],

[2] depends on the array configuration, system bandwidth, SNR, etc

however, for simplicity, in the following equations equal variance for all estimates is assumed.

The functional relationship between the target state and observation represented in (3.2) does not necessarily need to be Gaussian pdf, although it is very commonly used. To clarify, a likelihood function does not represent the estimate of a target state, it calculates the probability of obtaining the observation $\mathbf{z}_t$ provided the target is in state ζ. In most cases likelihood functions are not pdfs on the target state space and do not integrate to one over the target state space.

If the normalized residual of the k-th estimate for a target at location ζ is defined as

$$u_\zeta^k = \frac{h_{s,j}(\zeta) - \mathbf{z}_{t,k}}{\sqrt{R_t{}^{[sj]}}} \tag{3.3}$$

the likelihood function can be rewritten as

$$p(Z_t{}^{[sj]}|\zeta) = \prod_{k=1}^{K} \frac{exp(-\frac{1}{2}(u_\zeta^k)^2)}{\sqrt{2\pi R_t{}^{[sj]}}}. \tag{3.4}$$

The maximum likelihood estimation can be generalized by a maximum-likelihood-type estimator as

$$p(Z_t{}^{[sj]}|\zeta) = \prod_{k=1}^{K} \frac{exp(-p(u_\zeta^k))}{\sqrt{2\pi R_t{}^{[sj]}}} \tag{3.5}$$

where p is a symmetric, positive-definite function with a unique minimum at zero. With $p(u) = \frac{u^2}{2}$ the likelihood function is as in Eq. (3.4). In [118] the Welsh function, $p(u) = \frac{c_\alpha^2}{2}(1 - exp(-(u/c_\alpha)^2))$, is used for outlier suppression. The tuning constant c_α used in the Welsh function is used to make the maximum likelihood estimator more robust or more efficient. Other functions are also compared such as for example a least power function with $p = 1.2$, Cauchy function with 95 % efficiency or a German Macalure function.

When more information is available such as the probability of detection of the sensor, $p_D(\zeta)$, as a function, the two likelihood functions can be combined by pointwise multiplication. Incorporating the model for probability of detection of the sensor and the model for the false alarm intensity, $\varrho_F(\mathbf{z}_{t,k})$, the likelihood function for maximum one detection per target can be given by

$$p(Z_t{}^{[sj]}|\zeta) \propto \frac{p_D(\zeta)}{\varrho_F(\mathbf{z}_t)} p\left(\mathbf{z}_t|\zeta\right) + \left(1 - p_D(\zeta)\right). \tag{3.6}$$

Any information that can be put into the form of a likelihood function can be integrated in the overall likelihood function by combination, including subjective information.

The extension of the single target likelihood to a multiple target likelihood is straightforward. Typically an association variable, β, describing the measurement to target association is used. β can be defined as a vector with $\beta(i)$ being the index of the observation associated with target i and $\beta(i) = 0$ describes the event of no detection. The likelihood is then given by:

$$\begin{aligned} &p(Z_t{}^{[sj]}|\zeta^{(1)}, ..., \zeta^{(N)}) \\ &\propto \sum_{\beta} \left(\prod_{j:\beta(j)\neq 0} \left[\frac{p_D(\zeta^{(j)})}{\varrho_F\left(\mathbf{z}_t{}^{(\beta(j))}\right)} p\left(\mathbf{z}_t{}^{(\beta(j))}|\zeta^{(j)}\right) \right] + \prod_{j:\beta(j)=0} \left[1 - p_D(\zeta^{(j)})\right] \right). \end{aligned} \tag{3.7}$$

The likelihood function presented above considers only measurements by one transmitter-receiver pair. For the scenario used here, the above likelihood function should be generalized to consider the estimates of all transmitter-receiver pairs. The generalization is as follows:

$$p(Z_t|\zeta) = \prod_{s=1}^{N_s} \prod_{j=1}^{N_{sj}} p(Z_t{}^{[sj]}|\zeta) \tag{3.8}$$

where N_s is the number of sensors in the network and N_{sj} is the number of receivers on each sensor. Considering each transmitter-receiver pair as a single sensor, the joint likelihood function is defined as the product of all contributing single sensor likelihood functions.

The multiple sensor multiple target likelihood function described above is a generalised likelihood which can be used in scenarios where possibly different sensors can be used, some of which provide multiple estimation parameters per sensor. In this case the measurement function $h_{s,j}(\zeta)$ and the outlier suppression function $p(u)$ should be chosen appropriately to represent the provided sensor estimates.

3.4.3 Comparison of deterministic and statistical localisation

Previously on single sensor platform target range estimates are 'fused' to estimate a target location by ellipse intersection. This method does not take into consideration the variance of the target observations, and instead calculates a point intersection which deterministically represents the possible target location. A noisy estimated parameter can thus result in a wrong target position estimation for which there is no additional information on its accuracy. When multiple sensors are used, multiple location estimates of the same target are obtained. The mean of this cluster of estimates can then represent the final estimate of the target by the system of multiple sensors. Another method to fuse range estimates is by using a likelihood function as derived in Section 3.4.2. This function maps target observations to the target state space. Different noise distributions can be considered and used. The likelihood function defines the probability that a given observation is generated by a certain state from the state space. Thus the mapping in state space results in peaks in the likelihood function where the most likely target state is. In addition to target observations, sensor properties and context information may also be incorporated within the likelihood function. In case of multiple sensors, the likelihood functions with respect to each sensor may be multiplied and the product represents

a joint likelihood function that ideally describes the position of the target in the scenario more accurately compared to single sensor. A target state is usually extracted by searching for the peaks in the likelihood function.

In Fig. 3.11a the likelihood function over the scenario is represented for a target whose position is depicted as a green point in the scenario. The range estimates by the sensor are noisy and represented by the ellipses in blue. The sensor location (transmitter and two receivers on each side) are represented as black diamonds. If the ellipses are directly used to calculate the target location, the target location estimate is the one depicted by the blue point. As can be seen in the figure, this location estimate is not identical with the actual target position, i.e. the estimate is noisy. In this specific scenario example the peak of the likelihood function (dark red area) would not provide a more accurate location estimate. However the likelihood of the real target location is not zero as would be the case with the deterministic ellipse intersection. If however two sensors are used as depicted in Fig. 3.11b, the difference in the accuracy of the estimates is significant. The likelihood function in Fig. 3.11b has two noticeable peaks one of which is more pronounced.

In multiple target scenarios the ellipse intersections are not unique and the observation-target association becomes more cumbersome. As can be seen in Fig. 3.12a for the single sensor case and in Fig. 3.12b when two sensors are used, the likelihood function for localisation of the two targets has peaks close to the actual target positions.

For the derivation of the likelihood function shown in these figures no explicit observation-target association has been used and thus has multiple peaks which may be ambiguous. In single target scenarios the highest peak can be chosen as an estimate of the target state, however in scenarios where the number of targets is not known the question arises as to how many peaks should be extracted and how. Incorporation of observation-target association as in Eq. (3.7) may resolve some of the ambiguities. Additional information due to the target propagation over time would also resolve some of the ambiguities and is considered in Chapter 4.

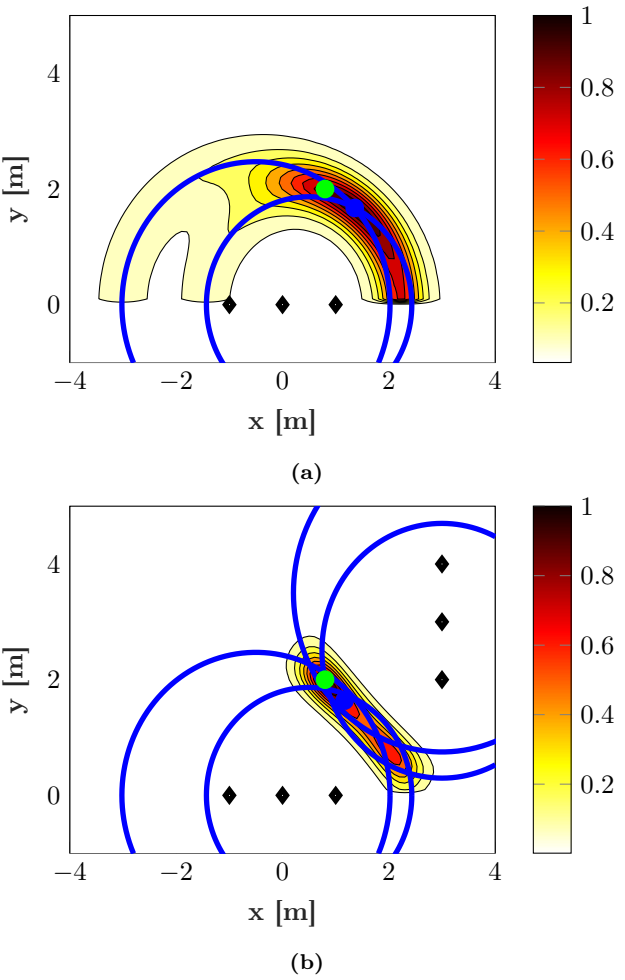
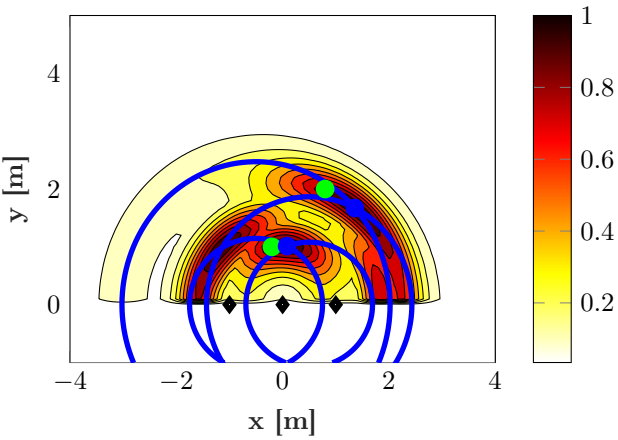
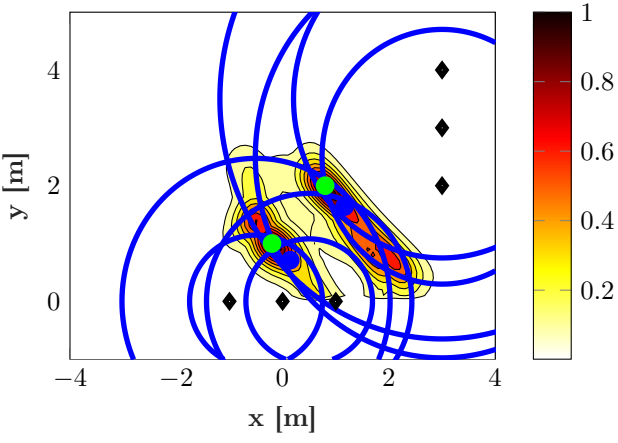


Figure 3.11: Sensor likelihood function overlaid by blue ellipses representing noisy target range estimates for a) single sensor and b) multiple sensors. Target location is depicted as green point and its estimate using ellipse intersection is depicted as blue point.



(a)



(b)

Figure 3.12: Sensor likelihood function overlaid by blue ellipses representing noisy target range estimates for a) single sensor and b) multiple sensors. Location of targets are depicted as green points and their estimates using ellipse intersection are depicted as blue points.

3.4.4 Likelihood function comparison based on experimental data

The joint likelihood function defined in Section 3.4.2 is used on a two person scenario surrounded by three UWB sensor nodes as described in Fig. 3.6. In all figures below the actual target positions are represented by the black and magenta points. The ground truth for actual positions of the targets could not be accurately measured and is thus an approximation itself.

In Fig. 3.13 a sample of the derived likelihood function is shown using the range estimates as used for the centralised and distributed fusion methods i.e. using all range estimates from all sensors and using only the range estimates that can be used to locally estimate a target location. In this specific snapshot it seems that using the distributed method is better since most of the clutter is filtered out and thus can not be mistaken as an additional target as in the centralised approach (see Fig. 3.13a upper right corner of the scenario).

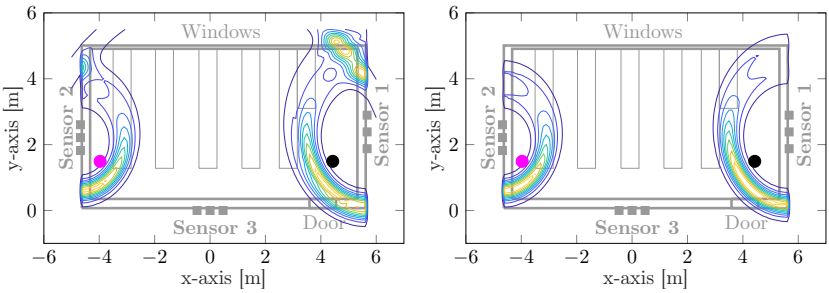


Figure 3.13: M-type likelihood function sample comparing data selection influence using a) all range estimates from the sensors and b) selected range estimates as used in the distributed fusion method

However, now taking another sample of the scenario (Fig. 3.14), it can be seen that by using the distributed approach one of the targets is not

represented in the likelihood. The reason for this is that the single detection of the second target is filtered out and discarded in the distributed approach. In the centralised approach (Fig. 3.14a), the target is represented in the likelihood but a peak can not be accurately extracted.

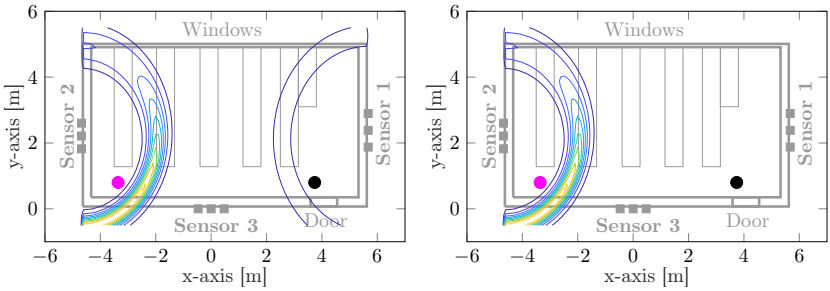


Figure 3.14: M-type likelihood function sample comparing data selection influence using a) all range estimates as used in the centralised fusion method and b) selected range estimates as used in the distributed fusion method

In Fig. 3.15 it can be noticed that both targets are properly represented in the likelihood function. The same time step as in Fig. 3.14 is depicted, however range tracking has been applied on the range estimates of each sensor. Due to the range tracking even though there are no target measurements provided by one of the transmitter-receiver pairs of the sensor, target estimates which can be used for localisation of the second target are available.

In Fig. 3.16 a sample of the likelihood function is shown where all range estimates from all sensors (top left), selected range estimates as in the distributed approach (top right), all range estimates after range tracking (bottom left) and selected range estimates after range tracking (bottom right) are used. What can be seen here is that range tracking removes some of the clutter detections without using the distributed approach for filtering out possible clutter. In addition, when range tracking is used, in both cases

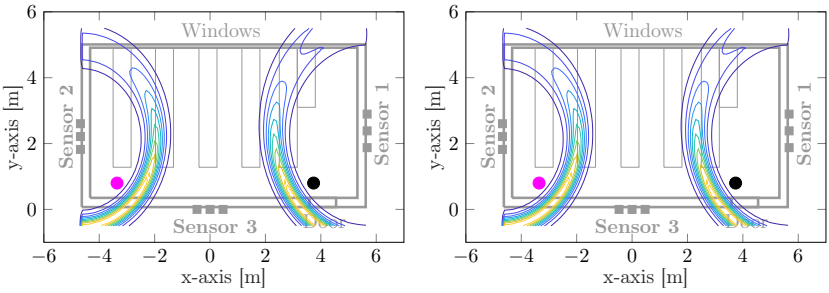


Figure 3.15: M-type likelihood function sample comparing data selection influence using a) centralised approach with tracked ranges and b) distributed approach with tracked ranges

the likelihood represents both targets, however in the centralised fusion method, the target states can be selected in an easier manner as peaks from the likelihood function. The peaks of the likelihood function when using the distributed approach are much wider.

The likelihood function for the different algorithm choice has been applied on many different scenarios with different sensor positions. It is difficult to derive a conclusion on which algorithm choice is better or worse. Applying range tracking filters out a lot of the clutter present in the range estimates of the different transmitter receiver pairs. In addition it also provides better target separation. However it also requires additional processing on the sensor side which can lead to delays in providing range estimates to the fusion center.

Using the distributed approach clutter is also removed by assuming that any target must be detected by both transmitter receiver pairs of a sensor. However some target information is also lost, making the estimation of the target position less accurate and more difficult. One important thing which can be noticed from our investigation is that in areas where a target can

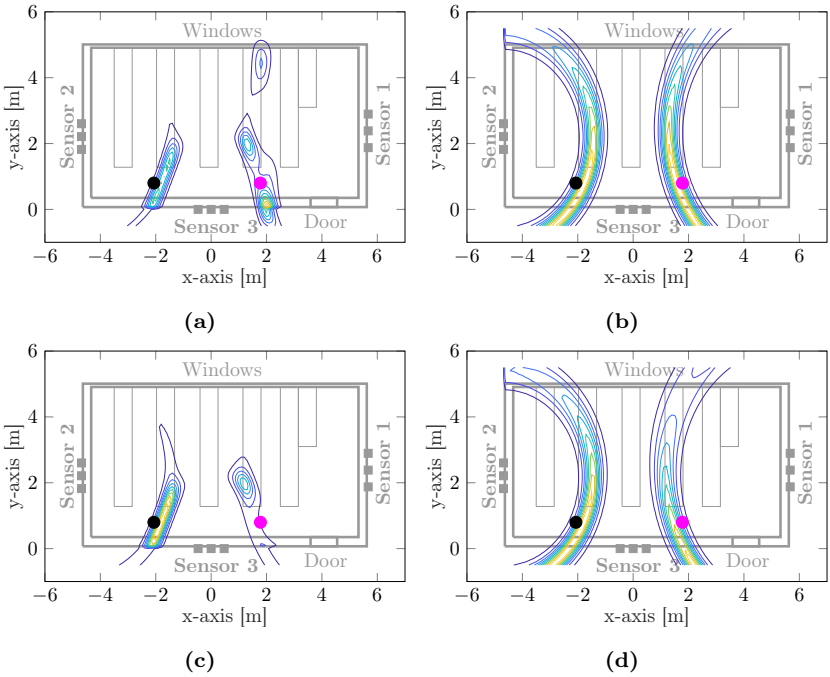


Figure 3.16: M-type likelihood function sample comparing data selection influence using a) all range estimates as used in the centralised fusion method b) selected range estimates as used in the distributed fusion method c) all range estimates after range tracking and d) selected range estimates after range tracking

be detected by more than one sensor, the peak of the likelihood function is more concentrated and easier to estimate. This is mainly accounted to the fact that the targets are observed from different aspect angles and data from transmitter-receiver pairs which are not too close to each other is used.

The aim of this thesis is to define a framework for person localisation. Range tracking seems to have a positive impact on filtering out clutter and improving target detections by filling in the gaps of target missdetection,

however it may be too expensive computationally to be done on the sensor side, thus a framework for target localisation and tracking in the fusion center is defined. Both centralised and distributed fusion approaches are further investigated, however range tracking is not used. Instead the tracking is done simultaneously with the localisation. Further details are given in Chapter 4.

3.5 Concluding remarks

In this chapter a distributed UWB sensor network is described. The main purpose of using a distributed sensor network is extension of the coverage area. Another advantage it offers is improvement in the system localisation capabilities in the overlap area. The processing of the measured impulse responses is also distributed in the network for reduction of the data transmission load between the network components.

As part of the work in this thesis a distributed sensor network concept is developed and implemented as demonstrator. The system, sensor and fusion center framework and architecture are described in this chapter. Different sensor cooperation and operating principles are investigated. Based on the sensor constellation and possibilities for synchronisation between the transmitters and receivers on the platforms, possibilities are explored for TOA, TDOA and hybrid TOA/TDOA localisation.

Fusing data from multiple sensors may be tricky due to possible additive noise in the sensor estimates as well as due to missdetections. Typically target observations by different sensors need to be associated to a target, especially in the regions where multiple sensors detect the targets. In this chapter the typical multiple sensor fusion challenges are described. Additionally a multiple sensor likelihood function is derived for the TOA based localisation case.

As can be seen in the likelihood function samples shown in this chapter the use of multiple distributed sensors for multiple person localisation increases the precision with which targets can be localised in overlapping areas.

CHAPTER 4

MULTIPLE PERSON TRACKING IN DISTRIBUTED UWB NETWORK

In the previous chapter a distributed sensor network is introduced for improvement of the detection and localisation of multiple targets. It was also shown how challenging combining observations of multiple sensors can be depending on the scenario geometry. The likelihood function presented in Section 3.4.2 presents a method for linking the observations of the sensors to the target state. The instantaneous likelihood only takes into consideration current target observations by the sensors. However due to the many missed detections by some sensors, it is worth to consider the information due to the target propagation and previous target detections by the sensors for improving the target detectability and localisation performance.

In this chapter introduction into statistical sensor data fusion and tracking is given with a brief review of possible methods and detailed explanation of the PHD filter as a chosen algorithm to further improve the multiple target localisation methodology presented in this thesis. The developed multiple target multiple sensor localisation and tracking method is applied on the same scenario in two ways: by using local target location estimates from each sensor as shown in Chapter 2, and by using all target range estimates from all synchronised transmitter-receiver pairs of the sensor network as suggested in Chapter 3.

4.1 Introduction

According to the revised definition of the Joint Directors of Laboratories (JDL) Data Fusion Model [120],

> Data fusion is the process of combining data or information to estimate or predict entity states.

Data is combined to estimate or predict the state of some aspect of the world. The objective is often estimation of the physical state of a system: identity, location, track over time, etc. of existing targets. Typically the input of a sensor data fusion system consists of the following:

1. data gathered/observed by sensors
2. data and command inputs by human operators or users
3. a-priori data from a pre-established database

The output of a data fusion system is intended to support human decision processes and can be provided at different level of inference, that is for instance parametric data for identification, time and local geometric analysis for entity behaviour inference, situation/goal analysis for situation assessment inference and contextual analysis for threat analysis inference.

Target tracking can be summarised as an estimation of the number of targets and their states at each point in time using a set of noisy measurements and previous target states. A state contains all the relevant information to describe the target such as location, velocity, identity, etc. Within this work targets are localised in 2D, thus four-term target states containing the x and y coordinates of the target and the respective velocities are estimated. Target identity is not considered. Nevertheless, the methods explained here are expandable for differently defined target states.

Within this chapter and thesis, UWB sensors are used to provide target observations (range data) and a-priori information on the possible scenario extent is used. There is no a-priori information from a established database since for the application scenarios considered there would be no option of

gathering a-priori information and instead the system would have to run in real-time and adapt on the available scenario challenges in real time.

4.2 Bayesian single target tracking

The Bayes filter is a successful approach to sequential estimation of the state of a dynamic system by using a sequence of noisy measurements. It is based on the theory of probabilistic filtering presented in [121]. A tracking system is considered Bayesian if there is a prior distribution on the state of the targets with probabilistic description of the target's motion characteristics, a likelihood function to define the relationship between the sensor observations and the target state, and a posterior distribution on the set of target states (joint distribution) defining the output of the tracker. It requires a good dynamic model that describes the state evolution over time as accurate as possible and a measurement model that relates the noisy measurements to the state.

The main idea of the single-target Bayes filter is to construct a posterior pdf $p(\mathbf{x}_t|Z_{1:t})$ of a state $\mathbf{x}_t$ based on all the available information, including the sequence of received measurements $Z_{1:t} = \{Z_1, Z_2, ...Z_t\}$ up to time t. Recursive filtering provides an estimate sequentially every time a new measurement is available. It incorporates two stages: prediction and update.

In the prediction stage a system model is used to predict the pdf of the state from time $t-1$ to a given time instant t:

$$p\left(\mathbf{x}_t|Z_{1:t-1}\right) = \int p\left(\mathbf{x}_t|\mathbf{x}_{t-1}\right) p\left(\mathbf{x}_{t-1}|Z_{1:t-1}\right) d\mathbf{x}_{t-1} \tag{4.1}$$

where $p\left(\mathbf{x}_t|\mathbf{x}_{t-1}\right)$ is the state transition pdf defining the target motion. Equation 4.1 is also known as the Chapman-Kolmogorov equation.

In the update stage the measurements received at time t are used to

correct or modify the predicted pdf by using the Bayes' formula

$$p\left(\mathbf{x}_t|Z_{1:t}\right) = p\left(\mathbf{x}_t|Z_t, Z_{1:t-1}\right) = \frac{p\left(Z_t|\mathbf{x}_t\right) p\left(\mathbf{x}_t|Z_{1:t-1}\right)}{\int p\left(Z_t|\zeta\right) p\left(\zeta|Z_{1:t-1}\right) d\zeta}. \tag{4.2}$$

with $p\left(Z_t|\mathbf{x}_t\right)$ being the measurement likelihood function. From the posterior pdf the optimal state estimates with respect to a given criterion can be computed. In the case of minimum mean squared error (MMSE) the optimal state estimate is $\hat{\mathbf{x}_t}^{MMSE} = E\left\{\mathbf{x}_t|Z_{1:t}\right\} = \int \mathbf{x}_t p\left(\mathbf{x}_t|Z_{1:t}\right) d\mathbf{x}_t$, whereas for MAP it is $\hat{\mathbf{x}_t}^{MAP} = \underset{\mathbf{x}_t}{\operatorname{argmax}}\, p\left(\mathbf{x}_t|Z_{1:t}\right)$.

In general the solution can not be computed analytically. An analytic solution exists only in some situations such as the KF in linear-Gaussian cases. In other cases suboptimal solutions such as EKF or UKF or particle filters are used.

4.2.1 The Kalman filter

The KF assumes perfect measurement to track association and linear-Gaussian density for representing the measurement likelihood and track evolution model. The track evolution model, $p\left(\mathbf{x}_t|\mathbf{x}_{t-1}\right)$ describes the target dynamics from time $t-1$ to time t. In the linear-Gaussian case it is given by:

$$p\left(\mathbf{x}_t|\mathbf{x}_{t-1}\right) = \mathcal{N}\left(\mathbf{x}_t; F_t\mathbf{x}_{t-1}, Q_t\right) \tag{4.3}$$

where F_t denotes the linear state transition over time (state evolution matrix) and Q_t is the process noise covariance describing the uncertainty in state evolution.

The measurement likelihood function $p\left(\mathbf{z}_t|\mathbf{x}_t\right)$ models the sensor characteristics and corresponds to $p\left(Z_t|\mathbf{x}_t\right)$ for single observation per sensor at a given time. Its covariance describes the measurement quality. In the

linear-Gaussian case it is given by

$$p\left(\mathbf{z}_t|\mathbf{x}_t\right) = \mathcal{N}\left(\mathbf{z}_t; H_t\mathbf{x}_t, R_t\right). \tag{4.4}$$

where H_t is the observation matrix describing the state contribution to the observation and R_t is the observation noise covariance (uncertainty in the observation). The process and observation noise are Gaussian, uncorrelated and zero-mean.

The optimal estimation density $p\left(\mathbf{x}_t|Z_{1:t}\right)$ is given by

$$p\left(\mathbf{x}_t|Z_{1:t}\right) = \mathcal{N}\left(\mathbf{x}_t; m_t, P_t\right). \tag{4.5}$$

The KF filter recursive equations are given as follows. Given a posterior at time $t-1$,

$$p\left(\mathbf{x}_{t-1}|Z_{1:t-1}\right) = \mathcal{N}\left(\mathbf{x}_{t-1}; m_{t-1}, P_{t-1}\right) \tag{4.6}$$

the prediction step calculates the prior as in Eq. (4.1) and results in

$$p\left(\mathbf{x}_t|Z_{1:t-1}\right) = \mathcal{N}\left(\mathbf{x}_t; m_{t|t-1}, P_{t|t-1}\right) \tag{4.7}$$

with

$$\begin{aligned} m_{t|t-1} &= F_t m_{t-1} \\ P_{t|t-1} &= F_t P_{t-1} {F_t}^T + Q_t. \end{aligned} \tag{4.8}$$

Using the prior, the posterior density is then calculated as in Eq. (4.2) and defined as in Eq. (4.5) with

$$\begin{aligned} m_t &= m_{t|t-1} + K_t\left(\mathbf{z}_t - H_t m_{t|t-1}\right) \\ P_t &= \left(I - K_t H_t\right) P_{t|t-1} \\ K_t &= P_{t|t-1} {H_t}^T \left(H_t P_{t|t-1} {H_t}^T + R_t\right)^{-1}. \end{aligned} \tag{4.9}$$

The Kalman gain, K_t, has a high value if the predicted covariance is large or the innovation covariance, $H_t P_{t|t-1} {H_t}^T + R_t$, is small. In this case the innovation, $\mathbf{z}_t - H_t m_{t|t-1}$, is reliable and the target estimate relies more on the innovation then on the prediction. If on the other hand the Kalman gain is low the filter follows the model predictions more closely and has low confidence in the sensor observations.

If the process and observation noise are not Gaussian, but symmetric and with finite moments, the KF would still provide good estimates. However if the noise distribution is skewed, the KF would provide misleading results.

4.2.2 The extended Kalman filter

The EKF is a variant of the KF that can be used when the state and/or observation model are non-linear. The process and observation noise are again Gaussian, uncorrelated and zero-mean. The process model is now described by a non-linear state transition function, $f(\mathbf{x}_t)$, and the observation model is also a non-linear function, $h(\mathbf{x}_t)$, describing the matching of the target states to the sensor observations. Both function should be differentiable.

The EKF follows closely the KF derivation with an additional step of linearisation of the process and observation models using Taylor series, i.e.

$$\begin{aligned} f(\mathbf{x}_t) &\approx f(m_{\mathbf{x}_t}) + J_F(m_{\mathbf{x}_t})(\mathbf{x}_t - m_{\mathbf{x}_t}) \\ h(\mathbf{x}_t) &\approx h(m_{\mathbf{x}_t}) + J_H(m_{\mathbf{x}_t})(\mathbf{x}_t - m_{\mathbf{x}_t}) \end{aligned} \tag{4.10}$$

where $\mathbf{x}_t \sim \mathcal{N}(m_{\mathbf{x}_t}, P_{\mathbf{x}_t})$ and J_F and J_H are the Jacobian matrices of f and h respectively.

The first terms in the linearisation contribute to the approximate means of the process and observation model i.e.

$$\begin{aligned} &E[f(\mathbf{x}_t)] \approx f(m_t) \quad \text{and} \\ &E[h(\mathbf{x}_t)] \approx h(m_t). \end{aligned} \tag{4.11}$$

The second terms define the approximate covariances of the respective functions i.e.

$$\begin{aligned} Cov\left[f(\mathbf{x}_t)\right] &= E\left[\left(f(\mathbf{x}_t) - E\left[f(\mathbf{x}_t)\right]\right)\left(f(\mathbf{x}_t) - E\left[f(\mathbf{x}_t)\right]\right)^T\right] \\ &\approx E\left[\left(f(\mathbf{x}_t) - f(m_t)\right)\left(f(\mathbf{x}_t) - f(m_t)\right)^T\right] \\ &\approx J_F(m_t) P_{\mathbf{x}_t} J_F^T(m_t) \quad \text{and} \end{aligned} \tag{4.12}$$

$$\begin{aligned} Cov\left[h(\mathbf{x}_t)\right] &= E\left[\left(h(\mathbf{x}_t) - E\left[h(\mathbf{x}_t)\right]\right)\left(h(\mathbf{x}_t) - E\left[h(\mathbf{x}_t)\right]\right)^T\right] \\ &\approx E\left[\left(h(\mathbf{x}_t) - h(m_t)\right)\left(h(\mathbf{x}_t) - h(m_t)\right)^T\right] \\ &\approx J_H(m_t) P_{\mathbf{x}_t} J_H^T(m_t) \end{aligned} \quad . \tag{4.13}$$

The Jacobian matrix of the models is used in the equations for state covariance and gain calculation. The approximate predicted and posterior pdf are approximated by a Gaussian distribution as in Eq. (4.7) and (4.5).

The EKF equations can thus be summarised as the following prediction equations:

$$\begin{aligned} m_{t|t-1} &= f(m_{t-1}) \\ P_{t|t-1} &= J_F(m_{t-1}) P_{t-1} J_F^T(m_{t-1}) + Q_{t-1} \end{aligned} \tag{4.14}$$

and the following update equations:

$$\begin{aligned} m_t &= m_{t|t-1} + K_t\left(\mathbf{z}_t - h(m_{t|t-1})\right) \\ P_t &= \left(I - K_t J_H(m_{t|t-1})\right) P_{t|t-1} \\ K_t &= P_{t|t-1} J_H^T(m_{t|t-1})\left(J_H(m_{t|t-1}) P_{t|t-1} J_H^T(m_{t|t-1}) + R_t\right)^{-1}. \end{aligned} \tag{4.15}$$

Since the Jacobian of the process and observation model are dependent on both state and time, they are not constant and thus need to be computed

online at each time step as the recursions are calculated. This makes the EKF more computationally demanding compared to the KF. Additionally, it is very important to initialise the filter accurately since in the recursion equations it uses linearisation of the model functions over the estimated or previous target states. It has to be taken care that the state predictions are close to the true state so that the second order terms of the Taylor series expansion are indeed insignificant and only the first term i.e. the Jacobian can be used. In case of high non-linearity the filter may become unstable and diverge from the estimation of the true states. Methods using higher order Taylor expansion have been considered but are limited by their increasing complexity [49].

4.2.3 The unscented Kalman filter

The UKF is another sub-optimal approximation of the single-target Bayes filter. The basic idea is to approximate the probability distribution instead of the non-linear model function. It is more accurate than the EKF for estimating the spread of a random variable [122]. It can deal with severe non-linearities with computational complexity of the same order. As in the other versions of the KF, Gaussian distribution is assumed. The Gaussian density is approximated by a set of deterministically chosen sample points, known as sigma points. The sigma points are then propagated using the non-linear process model function, and later used to re-approximate the mean and covariance of the Gaussian distribution. The Jacobian does not need to be explicitly calculated, which is an advantage and reduces complexity when the functions are complex or non-differentiable.

4.2.4 Particle filters

Another sub-optimal solution for the single-target Bayes filter is to approximate the densities using particles, i.e. N_p independent identically distributed samples [123–125]. In this way any non-linear model with non-Gaussian noise may be approximated. These sequential Monte Carlo (SMC)

based Bayes filter approximations are often popular for their flexibility, efficiency and simplicity.

Each particle has a likelihood weight representing the probability of that particle being sampled from the pdf. The particles and their weight are propagated over time to define the pdf at subsequent time steps. Weight collapse is often an issue in particle filters, however it can be avoided by using resampling techniques. In the resampling step particles with very low weight are replaced by new particles from the proximity of particles with higher weight. Closely related method is the bootstrap filter [124, 126].A Rao-Blackwell technique is often incorporated within the particle filter to reduce the size of the state space [125].

4.3 Bayesian multiple target tracking

Most of the methods above handle single target state i.e. estimate the behaviour of a target independent on the surrounding environment. In a multiple target scenario, the number of targets may change over time and the duration of target existence in the area of interest is unknown. The sensors observing the scenario do not always detect all existing targets, and the sensor observations are corrupted by spurious measurements that do not belong to any of the existing targets (clutter). In other words, the available sensor observations are indistinguishable and may belong to any of the present targets or clutter. Multiple target tracking involves joint estimation of the number and state of the targets from the available observations.

Traditional techniques split the multiple target problem in two separate methods: measurement to target association and target state estimation based on the association and single target tracking approaches, some of which are discussed in [49, 127–129]. One of the approaches for measurement to target association is the nearest neighbour method [50]. It associates the closest measurement to a target. The closeness of the measurement and target is defined in some statistical sense. Since multiple target tracks and measurements are available it happens that multiple measurements may

associate to a target, and a measurement may associate to more than one target track. The global nearest neighbour considers all possible associations and chooses the one that minimizes the sum of statistical distances between the targets and measurements [50].

4.3.1 Multiple hypothesis tracking

MHT is a technique that considers all possible associations at each time step and resolves the uncertainty in subsequent time steps when more measurements are available. Whenever an observation is obtained, multiple hypotheses are formed for associating the observation to the different targets, as well as hypotheses considering the observation a false alarm or belonging to a new target. These hypotheses are propagated over time, and any new observation spans multiple hypotheses from each of the already present hypotheses. This increases the complexity and computational cost of the method. In practice different techniques are used to reduce this complexity and truncate the number of spanning hypotheses at a given time such as gating (only observations within the gate region of an existing target may be associated to it), hypothesis pruning (any hypotheses with low probability of existence are discarded) and hypothesis merging (any hypotheses propagating very close to each other are merged to one).

The target states and covariances of each hypothesis may be updated using observations using the KF, EKF or UKF update equations. Additionally the hypothesis scores are also calculated and used to choose the best fitting hypothesis to define the target states. Practical implementations of the MHT are discussed in [50, 127].

4.3.2 FISST and RFS theory

The purpose of finite set theory (FISST) is to extend the idea of propagating the expectation of the posterior density from single target to multi-target problem. Some of the challenges are that the multi-target state and measurement have unknown dimension, changing in time, and at any time a

new target may appear or an existing target may disappear. In FISST terminology, the set X_t and the observation set Z_t are RFSs of the states and observations, respectively, at time t [53].

The multi-target state and measurement are the collection of the state and measurements of each target i.e. the state is the set $X_t = \{\mathbf{x}_t{}^i\}_{i=1}^{N_{\mathbf{x}_t}}$ as above and the multi-target measurement is the set $Z_t = \{\mathbf{z}_t{}^i\}_{i=1}^{N_{\mathbf{z}_t}}$ with variable number of targets $N_{\mathbf{x}_t}$ and number of measurements $N_{\mathbf{z}_t}$ over time.

The uncertainties in the multitarget state X_t and observation Z_t are described respectively by the RFSs $\Xi_t = S_t(X_{t-1}) \bigcup \Gamma_t$ and $\Psi_t = \Theta_t(X_t) \bigcup C_t$. The RFS Ξ encapsulates all aspects of multi-target motion such as the time-varying number of targets, individual target motion, target birth, target spawning, and target interactions. $S_t(X_{t-1})$ denotes the RFS of surviving targets from the previous time step and Γ_t describes the newly arising targets at time t. The RFS Ψ encapsulates all sensor characteristics such as measurement noise, sensor field of view (state dependent probability of detection), and false alarms. $\Theta_t(X_t)$ is the RFS modelling the measurements generated by the targets and C_t models the RFS of the clutter and false alarms. Each $\mathbf{z}_t{}^i$ is either a noisy observation of one of the targets or clutter. Each target state is represented by $\mathbf{x}_t{}^i$. The set-based approach allows for varying number of targets to appear and disappear without any particular order while avoiding explicit data association.

4.3.3 Multiple target Bayes filter

In case of multiple targets, the single-target Bayes filter can be extended to jointly estimate the states of all targets. The multi-target Bayes filter propagates the multi-target posterior density to find an optimal solution in terms of the minimum covariance. The recursion involves evaluation of multiple integrals on the state space. The size of the state space grows exponentially with the number of targets.

The states of all $N_{\mathbf{x}_t}$ targets at time t are represented by the set $X_t =$

$\{\mathbf{x}_t{}^i\}_{i=1}^{N_{\mathbf{x}_t}}$. The prediction and update equations are as follows:

$$p\left(X_t|Z_{1:t-1}\right) = \int p\left(X_t|X_{t-1}\right) p\left(X_{t-1}|Z_{1:t-1}\right) dX_{t-1} \tag{4.16}$$

$$p\left(X_t|Z_{1:t}\right) = \frac{p\left(Z_t|X_t\right) p\left(X_t|Z_{1:t-1}\right)}{\int p\left(Z_t|X\right) p\left(X|Z_{1:t-1}\right) dX} \tag{4.17}$$

where $p\left(X_t|Z_{1:t}\right)$ is the multi-target posterior density which is conditional on the measurements $Z_{1:t}$ up to time t. $p(X_t|X_{t-1})$ is the transition density and $p\left(Z_t|X_t\right)$ is the measurement multi-target likelihood function. The integral used in Equations 4.16 and 4.17 is not an ordinary integral but a set integral as explained in [53].

Since the multi-target Bayes filter is NP hard i.e. numerically intractable, it is too complex to be applied in practical scenarios. One of its approximations based on the Point process theory leads to the PHD filter [53–55].

4.4 The PHD filter

The Bayes filter propagates the posterior probability which in general case is not feasible. However propagation of its first-order statistical moment is feasible, as proposed by Mahler in [53]. This first-order statistical moment is also known as the intensity function or the PHD. Within this thesis both terms are used. It is defined as $v_t\left(\mathbf{x}_t\right) = \int \delta_{X_t}\left(\mathbf{x}_t\right) p\left(X_t|Z_{1:t}\right) \delta X_t$ where $\delta_{X_t}\left(\mathbf{x}_t\right) = \sum_{w \in X_t} \delta_w\left(\mathbf{x}_t\right)$. The PHD is a multi-modal distribution over the target space and each mode, or peak, represents a high probability of target presence. At a given time the target population is described by a set-valued state, thus it operates on single target state space and avoids the complexities arising from data association. A drawback of this superposition target space is the loss of target identities. Thus if the target identity is needed, target labels have to be obtained in a separate post processing step [130].

The PHD is not a pdf and its integration over a finite subset of the space

gives an estimated number of the targets in this subset. The propagation of the posterior intensity function v_t uses the following recursion:

$$v_{t|t-1}(\mathbf{x}_t) = \int \phi_{t|t-1}(\mathbf{x}_t, \mathbf{x}_{t-1})\, v_{t-1}(\mathbf{x}_{t-1})\, d\mathbf{x}_{t-1} + \gamma_t(\mathbf{x}_t) \tag{4.18}$$

$$v_t(\mathbf{x}_t) = [1 - p_D(\mathbf{x}_t)] v_{t|t-1}(\mathbf{x}_t) + \sum_{\mathbf{z}_t \in Z_t} \frac{\psi_{z,t}(\mathbf{x}_t)\, v_{t|t-1}(\mathbf{x}_t)}{\kappa_t(\mathbf{z}_t) + \int \psi_{z,t}(\zeta)\, v_{t|t-1}(\zeta)\, d\zeta}. \tag{4.19}$$

The transition density in (4.18) is defined as:

$$\phi_{t|t-1}(\mathbf{x}_t, \mathbf{x}_{t-1}) = p_S(\mathbf{x}_{t-1})\, p(\mathbf{x}_t|\mathbf{x}_{t-1}) + \beta_{t|t-1}(\mathbf{x}_t|\mathbf{x}_{t-1}) \tag{4.20}$$

where $p(\mathbf{x}_t|\mathbf{x}_{t-1})$ is the single target transition density, $p_S(\mathbf{x}_{t-1})$ is the probability of survival and $\beta_{t|t-1}(\mathbf{x}_t|\mathbf{x}_{t-1})$ is the PHD for spawned target birth. $\gamma_t(\mathbf{x}_t)$ is the PHD for spontaneous birth of new targets at time t. In the update equation (4.19), the first term, $[1 - p_D(\mathbf{x}_t)] v_{t|t-1}(\mathbf{x}_t)$, relates to the undetected targets with $p_D(\mathbf{x}_t)$ being the probability of target detection. It is considered that targets for which there are no sensor observations available at a given time would still be 'updated' by reducing their weight contribution in the joint posterior intensity. In this way targets unobserved for very short periods of time would still produce a state estimate since their contribution in the joint posterior intensity would still be high. If however such targets are not observed for prolonged period of time their contribution in the posterior intensity decreases and at some point they would no longer result in a state estimate.

The second term of the update equation (4.19) deals with targets observed by sensors. Namely,

$$\psi_{z,t}(\mathbf{x}_t) = p_D(\mathbf{x}_t)\, p(\mathbf{z}_t|\mathbf{x}_t) \tag{4.21}$$

where $p(\mathbf{z}_t|\mathbf{x}_t)$ is the single target likelihood function. The single target

likelihood function depends on the type of sensor used for observation and the measurement function defining the expected sensor observations. The clutter intensity is defined as

$$\kappa_t\left(\mathbf{z}_t\right) = \lambda_t q_t\left(\mathbf{z}_t\right), \tag{4.22}$$

where λ_t is the Poisson parameter defining the expected number of false alarms and $q_t\left(\mathbf{z}_t\right)$ is the clutter/false alarm probability distribution over the measurement space.

The cardinality, or the number of targets in the area of interest A is estimated as

$$N_{\mathbf{x}_t} = \int_A v_t\left(\zeta\right) d\zeta. \tag{4.23}$$

The main assumptions of the PHD filter are independence of the measurements generated by each target, the clutter is Poisson and independent from target-based measurements, and that the predicted multi-target RFS is Poisson. Since the PHD function requires the computation of integrals that in general do not have a closed form, it is approximated either by a set of particles (using the Particle filter) as in [55] or by GM components as in [54].

4.4.1 Particle PHD filter

The SMC implementation of the PHD filter is based on the random sample approximation of the PHD

$$v_t\left(\mathbf{x}_t\right) = \sum_{i=1}^{N_p} w_t^i \delta_{\tilde{\mathbf{x}_t}^i}\left(\mathbf{x}_t\right) \tag{4.24}$$

where $\delta_{\tilde{\mathbf{x}_t}^i}\left(\mathbf{x}_t\right)$ is the Dirac delta function and $\{w_t^i, \tilde{\mathbf{x}_t}^i\}_{i=1}^{N_p}$ is the weighted particle set describing the intensity. $\tilde{\mathbf{x}_t}^i$ are random samples of this intensity and their corresponding weight is w_t^i.

Most particles are drawn from the importance distribution (posterior intensity) and describe the target states. An additional number of particles are drawn from the birth intensity to represent possibly new born targets. The estimated number of targets is given by the total mass (summation of weights), $N_{\mathbf{x}_t} = \sum_{i=1}^{N_p} w_t^i$. A resampling step is applied to so that the number of particles does not grow exponentially.

4.4.2 Gaussian Mixture PHD filter

In this implementation of the PHD filter, it is approximated by weighted GMs which are projected to the next time step by using the prediction equation. Let the posterior intensity $v_{t-1}(\mathbf{x}_t)$ at time $t-1$ be approximated as

$$v_{t-1}(\mathbf{x}_t) = \sum_{i=1}^{J_{t-1}} w_{t-1}^i \mathcal{N}\left(\mathbf{x}_t; m_{t-1}^i, P_{t-1}^i\right) \tag{4.25}$$

where m_{t-1}^i and P_{t-1}^i are the mean and covariance of the J_{t-1} Gaussians representing the RFS of target states at time $t-1$.

The predicted intensity can then be represented as

$$v_{t|t-1}(\mathbf{x}_t) = v_{S,t|t-1}(\mathbf{x}_t) + \gamma_t(\mathbf{x}_t) \tag{4.26}$$

where the survival intensity that describes the intensity of the targets which survive from time $t-1$ is defined as

$$\begin{aligned} v_{S,t|t-1}(\mathbf{x}_t) &= \int \phi_{t|t-1}(\mathbf{x}_t, \mathbf{x}_{t-1})\, v_{t-1}(\mathbf{x}_{t-1})\, d\mathbf{x}_{t-1} \\ &= p_S \sum_{i=1}^{J_{t-1}} w_{t-1}^i \mathcal{N}\left(\mathbf{x}_t; m_{S,t|t-1}^i, P_{S,t|t-1}^i\right) \end{aligned} \tag{4.27}$$

when target spawning is not considered, i.e. $\beta_{t|t-1}(\mathbf{x}_t|\mathbf{x}_{t-1}) = 0$ and the probability of survival is state independent, i.e. $p_S(\mathbf{x}_t) = p_S$. The GM

parameters are defined as follows:

$$\begin{aligned} m^i_{S,t|t-1} &= F_t m^i_{t-1} \\ P^i_{S,t|t-1} &= F_t P^i_{t-1} {F_t}^T + Q_t \end{aligned}. \tag{4.28}$$

The birth intensity is defined as a sum of weighted GMs with J_{γ_t} components with weight $w^i_{\gamma_t}$, mean $m^i_{\gamma_t}$ and covariance $P^i_{\gamma_t}$:

$$\gamma_t(\mathbf{x}_t) = \sum_{i=1}^{J_{\gamma_t}} w^i_{\gamma_t} \mathcal{N}\left(\mathbf{x}_t; m^i_{\gamma_t}, P^i_{\gamma_t}\right). \tag{4.29}$$

The GM components used to define the birth intensity may be randomly distributed over the whole area of interest, or distributed only over specific areas of the scenario where possible new target births are expected. More on how the birth intensity may be defined is given in Section 4.5.1.

Now, let the predicted intensity $v_{t|t-1}(\mathbf{x}_t)$ be represented using $J_{t-1}+J_{\gamma_t}$ GMs:

$$v_{t|t-1}(\zeta) = \sum_{i=1}^{J_{t-1}+J_{\gamma_t}} w^i_{t|t-1} \mathcal{N}\left(\zeta; m^i_{t|t-1}, P^i_{t|t-1}\right). \tag{4.30}$$

The posterior intensity at time t is then

$$v_t(\mathbf{x}_t) = \left(1 - p_D(\mathbf{x}_t)\right) v_{t|t-1}(\mathbf{x}_t) + \sum_{\mathbf{z}_t \in Z_t} v_{D,t}(\mathbf{x}_t; \mathbf{z}_t) \tag{4.31}$$

where the first term deals with targets not detected by the sensors i.e. targets for which there is no observation to update their tracks. The $v_{D,t}(\mathbf{x}_t; \mathbf{z}_t)$ is the intensity of the detected targets and defined as:

$$v_{D,t}(\mathbf{x}_t; \mathbf{z}_t) = \sum_{i=1}^{J_{t-1}+J_{\gamma_t}} w^i_t(\mathbf{z}_t) \mathcal{N}\left(\mathbf{x}_t; m^i_t(\mathbf{z}_t), P^i_t(\mathbf{z}_t)\right). \tag{4.32}$$

The GM parameters are dependent on the target observations $\mathbf{z}_t$ and are defined as follows:

$$\begin{aligned} w_t^i(z) &= \frac{p_D(\mathbf{x}_t)\, w_{t|t-1}^i p\left(z|\mathbf{x}_t{}^i\right)}{\kappa_t(z) + \sum_{j=1}^{J_{t-1}+J_{\gamma_t}} p_D(\mathbf{x}_t{}^j)\, w_{t|t-1}^j p\left(z|\mathbf{x}_t{}^j\right)} \\ m_t^i(z) &= m_{t|t-1}^i + K_t^i\left(z - H_t m_{t|t-1}^i\right) \\ P_t^i(z) &= \left(I - K_t^i H_t\right) P_{t|t-1}^i \end{aligned} \tag{4.33}$$

with

$$K_t^i = P_{t|t-1}^i {H_t}^T \left(H_t P_{t|t-1}^i {H_t}^T + R_t\right)^{-1} \tag{4.34}$$

and the marginal likelihood function defined as

$$p\left(z|\mathbf{x}_t{}^i\right) = \mathcal{N}\left(z; H_t m_{t|t-1}^i, H_t P_{t|t-1}^i {H_t}^T + R_t\right) \tag{4.35}$$

4.4.3 Pruning and merging

When using the GMPHD, after the update step GMs with weight below a predefined threshold are pruned and GMs close to each other are merged. This is done to limit the number of GMs representing the posterior PHD to only relevant contributing mixtures. If the posterior PHD is represented as

$$v_t(\mathbf{x}_t) = \sum_{i=1}^{J_t} w_t^i \mathcal{N}\left(\mathbf{x}_t; m_t^i, P_t^i\right) \tag{4.36}$$

and the weights $w_t^1, ..., w_t^{J_\tau}$ are below the predefined threshold, τ, then the posterior PHD can be represented as

$$v_t(\mathbf{x}_t) = \frac{\sum_{l=1}^{J_t} w_t^l}{\sum_{j=J_\tau+1}^{J_t} w_t^j} \sum_{i=J_\tau+1}^{J_t} w_t^i \mathcal{N}\left(\mathbf{x}_t; m_t^i, P_t^i\right). \tag{4.37}$$

That is the new weights are determined as:

$$\overline{w}_t^i = w_t^i \frac{\sum_{l=1}^{J_t} w_t^l}{\sum_{j=J_\tau+1}^{J_t} w_t^j}. \tag{4.38}$$

The next step is to merge Gaussians whose distance between their means falls within a merging threshold, U. In the set of Gaussian components, $I = \{i = J_\tau+1, ..., J_t\}$, first the component with largest weight is chosen, i.e. $j = \underset{i \in I}{\operatorname{argmax}}\, w_t^i$. Then the set of components whose Mahalanobis distance to this component is below the predefined merging threshold is defined as

$$L = \{i \in I | \left(m_t^i - m_t^j\right)^T \left(P_t^i\right)^{-1} \left(m_t^i - m_t^j\right) \leq U\}. \tag{4.39}$$

The new Gaussian component for $l = 1$ is then defined as

$$\begin{aligned} \tilde{w}_t^l &= \sum_{i \in L} \overline{w}_t^i \\ \tilde{m}_t^l &= \frac{1}{\tilde{w}_t^l} \sum_{i \in L} \overline{w}_t^i m_t^i \\ \tilde{P}_t^l &= \frac{1}{\tilde{w}_t^l} \sum_{i \in L} \overline{w}_t^i \left(P_t^i + \left(\tilde{m}_t^l - m_t^i\right)\left(\tilde{m}_t^l - m_t^i\right)^T\right) \end{aligned} . \tag{4.40}$$

Next, the components in the set L are removed from the set I and the same procedure given above is repeated for $l = l + 1$ until the set I is empty.

Typically a maximum number of Gaussian components, J_{max}, is set for controlling the number of possible Gaussian components from growing exponentially and limiting the computational performance. In the case when $l > J_{max}$, the J_{max} components with largest weights are used.

4.4.4 Target state estimation

Gaussian terms with weight above a predefined confirmation threshold, T, are considered for target state estimation. The GM mean is considered to be the target state. The number of targets is considered to be equal to the number of Gaussians with weights above this threshold.

Thus the set of target estimates is

$$\widehat{X_t} = \{m_t^i : w_t^i > T\}. \tag{4.41}$$

4.4.5 Track labelling and identification

Identity declaration is a basic step of the processing chain of a data fusion system to associate a label to a target. This label can be semantic i.e. describing the class (for e.g. vehicle or human) of the target, or an identification number used for associating purposes. Pattern recognition techniques are employed to perform this task. Thus, this process is usually associated with the feature extraction approaches in order to extract relevant information from the data. Many different techniques may be employed in this stage such as artificial neural networks, support vector machines, Bayesian methods, clustering algorithms, etc. Within this work targets are not specifically identified since the PHD filter works specifically without data association. If the target identity is needed, a simple target labelling procedure as in [131] may be applied.

4.5 Extensions to the PHD filter

The PHD filter as described above in Section 4.4 has been extended using methods available in literature for application on the scenarios considered in this thesis. One extension is the use of measurement-based target birth for target track initialisation. The other extension considers the use of multiple sensors. The PHD filter assumes single observation per target, which is

not applicable when multiple sensors observe the target. Thus sequential update PHD filter is applied.

4.5.1 Target birth and initialisation

For having a working multiple target tracking algorithm based on the PHD filter, defining an appropriate birth intensity is very important. The intensity of the spontaneous birth RFS is usually defined as

$$\gamma_t(\mathbf{x}_t) = \sum_{i=1}^{N_{\gamma_t}} w_{\gamma_t}^i \mathcal{N}\left(\mathbf{x}_t; m_{\gamma_t}^i, P_{\gamma_t}^i\right) \tag{4.42}$$

where N_{γ_t} is the number of birth GMs at time t, $w_{\gamma_t}^i$ is the weight of the i-th GM with mean $m_{\gamma_t}^i$ and covariance $P_{\gamma_t}^i$.

In this formulation, the intensity function of the new-born targets is independent of the measurements, and generally it covers the entire state space since new targets may appear anywhere. This can also be seen in the formulation of the PHD filter equations, Eq. (4.18) and (4.19), where new targets are "born" in the prediction step. The mean of the newly born mixtures is chosen randomly from the state space, and the defining parameter is N_{γ_t}. Defining the birth intensity this way is inefficient. A simple alternative birth intensity definition is to use only one big Gaussian at the center of the field of view with large enough covariance to cover the whole field of view.

Measurement-based birth intensity as in [132] is more efficient for target initialisation compared to random initialisation. The subspaces where a target may appear are based on target observations and not randomly over the whole state space, i.e. newly born targets appear in regions with high likelihood. The weights of the Gaussians representing the new targets are given as $w_t^b = \widehat{n}_{t|t-1}/N_{\mathbf{z}_t}$, where $\widehat{n}_{t|t-1}$ is the expected number of newborn targets and $N_{\mathbf{z}_t}$ is the number of observations used to define the birth intensity. Only observations not previously used for updating persistent

targets are used for target initialisation. This means that at the initialisation step all available observations would be used. The mean of the Gaussians is directly taken from the provided observations if the observations and targets coordinate systems correspond, otherwise the observations are converted to the targets coordinate system. If incomplete observations are provided (as is the case in the work presented in this thesis), samples of the high likelihood regions of the likelihood function are used:

$$\gamma_t{}^{new}(\mathbf{x}_t) = \sum_{\mathbf{z}_t \in Z_t} \frac{p(\mathbf{z}_t|\mathbf{x}_t)\,\gamma_t(\mathbf{x}_t)}{\kappa_t(\mathbf{z}_t) + \int p(\mathbf{z}_t|\mathbf{x}_t)\,\gamma_t(\zeta)\,d\zeta}. \tag{4.43}$$

The birth intensity also depends on whether a target is persistent or not. Persistent targets are basically all existing targets in the state space. If a target is persistent it cannot be considered as a newborn target and a newborn target becomes persistent at the next time step. Thus the target state $\mathbf{x}_t$ can be extended to also contain a label which denotes if the target is persistent or newborn. The upgraded PHD filter equations are given in [132] with details for the SMC PHD filter.

Within this work measurement-based birth intensity is used for initialisation of new targets throughout the scenario as well as for initialisation of the PHD filter recursion.

4.5.2 Sequential-update PHD filter

Generalizations of the single-sensor PHD filter to a multiple sensor case have been originally proposed by Mahler in [53]. The update equation is however too complicated to be of practical use.

An inaccurate solution is often applied by fusing the observations of all sensors into one pseudo-sensor observation set and then applying the single sensor PHD filter. This solution is inaccurate since it violates the measurement model used to derive the update equation, i.e. single observation per target.

Different multiple sensor approximations are presented in [133]. The most common approach is to apply the single sensor update equation multiple times in succession as in [134]. The update equation of the multiple sensor PHD filter is then approximated as:

$$v_t\left(\mathbf{x}_t\right) = {v_{t|t-1}}^{[N_s]}\left(\mathbf{x}_t\right) \tag{4.44}$$

where N_s is the number of sensors used.

Initialising with

$${v_{t|t-1}}^{[0]}\left(\mathbf{x}_t\right) = v_{t|t-1}\left(\mathbf{x}_t\right) \tag{4.45}$$

and updating using the observations of each sensor sequentially where the update equation for the j-th sensor is defined as follows:

$$\begin{aligned} {v_{t|t-1}}^{[j]}\left(\mathbf{x}_t\right) = & \left[1 - {p_D}^{[j]}\left(\mathbf{x}_t\right)\right] {v_{t|t-1}}^{[j-1]}\left(\mathbf{x}_t\right) \\ & + \sum_{\mathbf{z}_t \in {Z_t}^{[j]}} \frac{{\psi_{z,t}}^{[j]}\left(\mathbf{x}_t\right) {v_{t|t-1}}^{[j]}\left(\mathbf{x}_t\right)}{{\kappa_t}^{[j]}\left(\mathbf{z}_t\right) + \int {\psi_{z,t}}^{[j]}\left(\zeta\right) {v_{t|t-1}}^{[j]}\left(\zeta\right) d\zeta} \end{aligned} \tag{4.46}$$

and

$${\psi_{z,t}}^{[j]}\left(\mathbf{x}_t\right) = {p_D}^{[j]}\left(\mathbf{x}_t\right) p\left(\mathbf{z}_t|\mathbf{x}_t\right) \tag{4.47}$$

where ${p_D}^{[j]}\left(\mathbf{x}_t\right)$ is the probability of detection of sensor j, ${Z_t}^{[j]}$ is the set of acquired measurements by sensor j at time t, ${\kappa_t}^{[j]}\left(z\right)$ is the clutter intensity for sensor j and $p\left(\mathbf{z}_t|\mathbf{x}_t\right)$ is the single target likelihood function for sensor j when $\mathbf{z}_t \in {Z_t}^{[j]}$.

The sequential PHD filter is highly dependent on the order the sensors are used. Thus if sensor 'reliability' is known or computed, the observations of the most 'reliable' sensors should be used first. To avoid the sensor order dependency, the sequential PHD filter can be applied over all possible sensor orders. This can be very expensive when many sensors are used. The updated intensities of the different sensor orders can then be merged. It has to be noted that before merging the intensity obtained from the different sensor orders a normalization with the number of possible sensor

order permutations should be applied.

For the GM implementation of the PHD filter, omitting spawning for simplicity, the number of GMs in $v_t(\mathbf{x}_t)$ is

$$J_t = (J_{t-1} + J_{\gamma_t}) \left(1 + |Z_t{}^{[1]}|\right) ... \left(1 + |Z_t{}^{[N_s]}|\right) \tag{4.48}$$

and it has an almost exponential growth with time and number of sensors. If $|Z_t{}^{[1]}| = ... = |Z_t{}^{[N_s]}|$ and J_{γ_t} is time independent, the growth is exponential. Thus the use of pruning and merging is very important.

4.6 Bistatic Range-Only Target Tracking Concepts

In range-only localization, to estimate the location of a target the only observation of the target considered is the target range. It is assumed that the location of the transmitters and receivers is known. The time needed for the signal to travel from the transmitter to a target and back to a receiver can be accurately measured and the distance from the transmitter to the target plus the distance from the target to the receiver represents the target bistatic range.

In this section two approaches for target localization and tracking based on range-only estimates are presented. Before the methods are elaborated, the common target state dynamics model is described. The first approach determines location estimates from the estimated ranges by an ellipse intersection method with simple local data association procedure to eliminate possible ghost target locations as described in section 2.6.2. Then a location tracking method is applied to smooth and merge the location estimates from the multiple sensors. This method is here also referred to as distributed fusion approach. The second approach is a direct tracking approach where the range estimates of all sensors are directly applied to determine the likelihood function used in the tracker as given in Section 3.4.2. The advantage of this method is that the range estimation errors are not directly

propagated into location estimation errors and target detections of a receiver is not discarded if the other receiver of the sensor did not detect the target. A block diagram of a multiple sensor target tracker is given in Fig. 4.1.

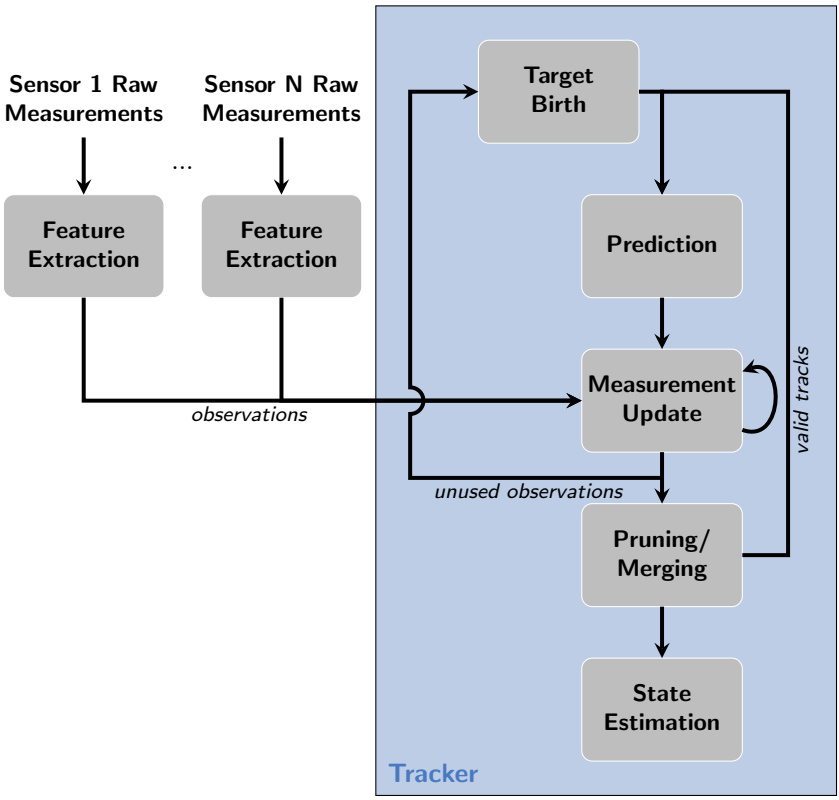


Figure 4.1: Block diagram of a multiple target tracker

The difference in the two approaches is in the observations provided by sensors to the fusion center. This makes the measurement functions in the measurement update step different, i.e. in the first case a linear function may be used and in the second case the non-linear relationship between

the target observation and target state is linearised. The specifics of the sensor models used for describing the measurement function are given in the description of each method below, i.e. Section 4.6.2 and 4.6.3 respectively. Due to the use of multiple sensors, the sequential update presented in Section 4.5.2 with the resective measurement function is applied. For the prediction stage of the tracker a target dynamic model is used as described in Section 4.6.1. For the target birth and target initialisation stage the measurement based model described in Section 4.5.1 is used and thus only unused measurements from previous processing step are used as input for the target initialisation. The pruning and merging stages are described in Section 4.4.3 and the finally the state estimation method is described in Section 4.4.4.

4.6.1 Target state dynamics

The target state is usually defined as a vector containing kinematic components such as target position, velocity and acceleration. Often constraints are used to control the growth of the specific components i.e. maximum speed for the velocity component. In this way the specific target motion can be better modelled. Other components such as for example target identity may also be part of the target state. The target state should be able to describe a target in the best possible way taking into consideration the specific scenario application. When defining the state space it should be taken into consideration that the target's motion is Markovian in the chosen state space and that the sensor likelihood functions depend only on the target state at the time of observation.

For the work presented in this thesis, the state vector of a target at time t is defined by the Cartesian (x, y) target position coordinates and velocity:

$$\mathbf{x}_t = [x_t, y_t, \dot{x}_t, \dot{y}_t]^T \tag{4.49}$$

Linear Gaussian target dynamics is considered:

$$p\left(\mathbf{x}_t|\mathbf{x}_{t-1}\right) = \mathcal{N}\left(\mathbf{x}_t; F_t\mathbf{x}_{t-1}, Q_t\right) \tag{4.50}$$

with

$$F_t = \begin{bmatrix} I_2 & dtI_2 \\ O_2 & I_2 \end{bmatrix} \text{ and } Q_t = \sigma_{\upsilon_t}^2 \begin{bmatrix} \frac{dt^4}{4}I_2 & \frac{dt^3}{2}I_2 \\ \frac{dt^3}{2}I_2 & dt^2 I_2 \end{bmatrix}$$

as the state transition matrix and the process noise covariance. dt is the time interval between two observations, $\sigma_{\upsilon_t}^2$ is the variance of the process noise $\upsilon_t \sim \mathcal{N}\left(0, Q_t\right)$, and I_n and O_n denote $n \times n$ identity and zero matrices respectively.

This model is also known as *constant-velocity*, or "nearly-constant-velocity" model [135] for non-manoeuvring targets. The process noise is a discrete-time vector-valued white noise and should only have a small effect on the target state accounting for unpredictable modelling errors due to slight manoeuvres, etc. The elements of υ_t correspond to noisy accelerations along the corresponding coordinates. It is assumed that υ_t is uncoupled across its components, i.e. uncoupled across x and y direction. By increasing the intensity of the process noise this model can be used for manoeuvring targets (*white-noise acceleration model [135]*).

This is a simple target dynamics model and target state. By modelling the process noise as a Wiener process the *constant-acceleration* or "nearly-constant-acceleration" model can be derived [135]. When using this model the target state needs to also include the target acceleration. This model assumes that the acceleration increment is independent over different sampling intervals, which is rarely true in practice. More on different dynamic models can be read in [136].

A more sophisticated interacting multiple model set may be used to cover different possibilities for target motion such as specific model for straight movement, model for coordinated turns etc [137]. In this case the target state should be extended to consider the angular turn at each time step.

Depending on scenarios and targets the target state and dynamic model should be upgraded to account for other target specific motion profiles i.e. if the desired target is a flying object the target state should be defined in 3D and maybe consider yaw, roll and pitch.

4.6.2 Location tracking in multiple sensor scenario

Since a target is typically detected by more than one sensor, these location estimates are fused to result in a single target location per target. The estimated target locations using this method contain significant amount of noise due to the propagation of range estimate errors into location estimate errors. In addition, pairs of closely spaced receiving antennas result in a flat intersection of two ellipses corrupting the final result significantly. These location estimates can be used as noisy location observations of the targets in the scenario and applied to a filter designed to fuse observations from multiple sensors to target location estimates for all targets in the scenario. In case of multiple targets, an association step of the observations to the different targets might be needed. Here a variation of the PHD filter is used. As mentioned earlier in this chapter the PHD filter can handle multiple target localisation without specific observation-target association. In Fig. 4.2 a block diagram of a target tracker based on this approach is presented.

The observation set at time t is defined as $Z_t = \bigcup_{s=1}^{N_s} {Z_t}^{[s]}$, where ${Z_t}^{[s]}$ is the set of location estimates determined using the data from sensor s at time t. N_s is the number of sensors used. The observation equation is defined as

$$\mathbf{z}_t{}^{[s]} = H_t \mathbf{x}_t + \omega_t{}^{[s]}, \tag{4.51}$$

where $H_t = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$ for a four-term target state, $\omega_t{}^{[s]} \sim \mathcal{N}\left(0, R_t{}^{[s]}\right)$ is the observation process noise and $\mathbf{z}_t{}^{[s]}$ is the observation vector at time t from sensor s.

In this method the observation $\mathbf{z}_t{}^{[s]}$ is not the directly estimated parameter

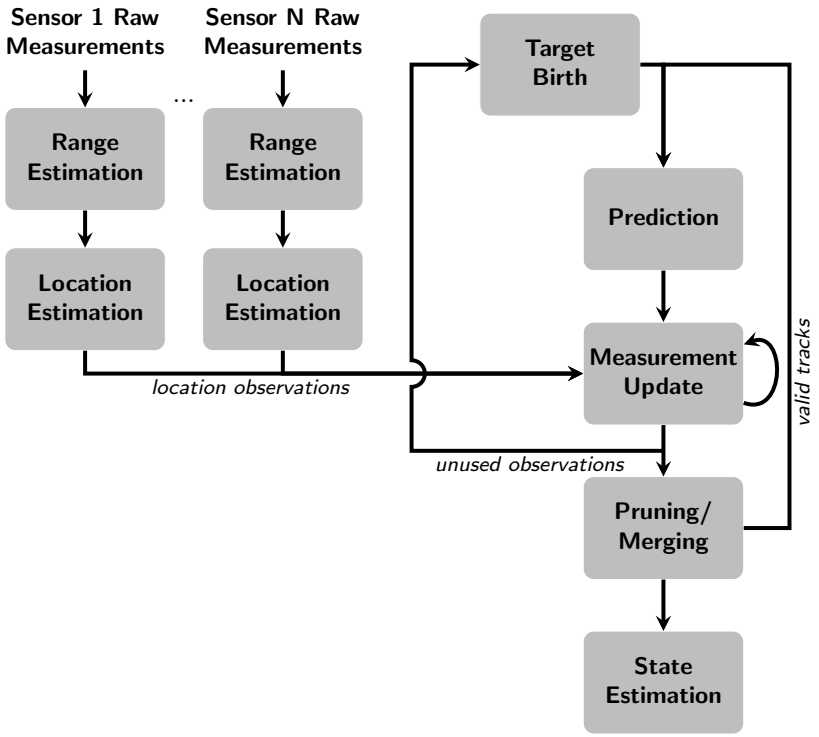


Figure 4.2: Block diagram of a tracker based on location tracking

from the received signal, but instead it is defined as the estimated location using the ellipse intersection method explained in Section 2.6.2 using the estimated ranges in the received signal by each of the receivers of the sensor. The observation process covariance is thus highly dependent on the range estimation error $\varepsilon_{s,j,t}$ from (2.17) of the two transmitter receiver pairs used for calculating the location estimate. The observation process covariance

can thus be defined as:

$$R_t{}^{[s]} = \left(\frac{\partial \tilde{r}_{s,1}(\mathbf{x}_t)}{\partial \mathbf{x}_t} \sigma_{s,1,t}^{-1} \left(\frac{\partial \tilde{r}_{s,1}(\mathbf{x}_t)}{\partial \mathbf{x}_t} \right)^T + \frac{\partial \tilde{r}_{s,2}(\mathbf{x}_t)}{\partial \mathbf{x}_t} \sigma_{s,2,t}^{-1} \left(\frac{\partial \tilde{r}_{s,2}(\mathbf{x}_t)}{\partial \mathbf{x}_t} \right)^T \right)^{-1} \tag{4.52}$$

where $\tilde{r}_{s,j}(\mathbf{x}_t)$ is the range equation for a target with state $\mathbf{x}_t$ with respect to receiver j of sensor s given in (2.17), $\frac{\partial \tilde{r}_{s,j}(\mathbf{x}_t)}{\partial \mathbf{x}_t}$ is the Jacobian of the range equation and $\sigma_{s,j,t}$ is the covariance of the range estimation error $\varepsilon_{s,j,t}$.

The classical single target likelihood is then

$$\begin{aligned} &p(\mathbf{z}_t{}^{[s]}|\mathbf{x}_t) \\ &\quad = \frac{1}{\sqrt{|R_t{}^{[s]}|(2\pi)^l}} \exp\left(-\frac{1}{2} \left(\mathbf{z}_t{}^{[s]} - H_t * \mathbf{x}_t \right)' R_t{}^{[s]\,-1} \left(\mathbf{z}_t{}^{[s]} - H_t * \mathbf{x}_t \right) \right) \end{aligned} \tag{4.53}$$

where l is the length of the measurement vector $\mathbf{z}_t{}^{[s]}$.

4.6.3 Direct range-to-location tracking in multiple sensor scenarios

For this approach the target observations used for the tracker are defined as the estimated ranges from all transmitter-receiver pairs of the observing systems. The concept is depicted in Fig. 4.3. The range of the targets with respect to each transmitter-receiver pair defines the observations at time t as $Z_t{}^{[s,j]} = \{\mathbf{z}_t{}^{[s,j]\,i}\}_{i=1}^{N_{\mathbf{z}_t}{}^{sj}}$ where $N_{\mathbf{z}_t}{}^{sj}$ is the number of targets detected by the j-th receiver of sensor s. The observation set is defined as $Z_t = \bigcup_{s=1}^{N_s} \bigcup_{j=1}^{N_{sj}} Z_t{}^{[s,j]}$ where N_s is the number of sensors used and N_{sj} is the number of receivers of sensor s.

At time instant t, the observation of a target with respect to receiver j is

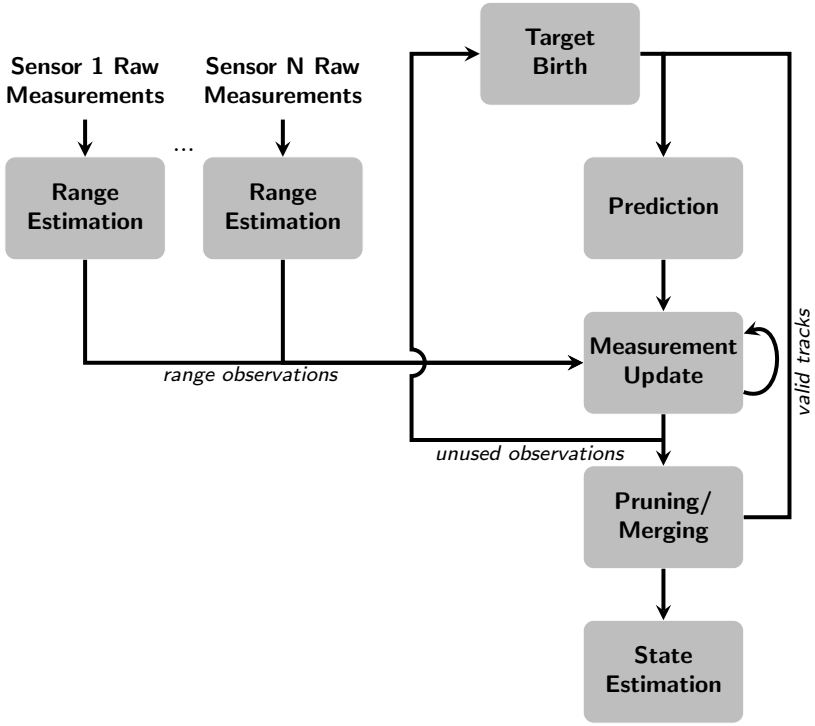


Figure 4.3: Block diagram of a tracker using all range estimates as observations

defined using the measurement equation as:

$$\mathbf{z}_t{}^{[s,j]} = h_{s,j}\left(\mathbf{x}_t\right) + \omega_t{}^{[sj]} \tag{4.54}$$

where

$$\begin{aligned} h_{s,j}\left(\mathbf{x}_t\right) &= r_{s,j}(\mathbf{x}_t) \\ &= \sqrt{\left(x_t - x_s\right)^2 + \left(y_t - y_s\right)^2} + \sqrt{\left(x_t - x_{s,j}\right)^2 + \left(y_t - y_{s,j}\right)^2} \end{aligned} \tag{4.55}$$

is the measurement equation, which is same as the range equation (2.16), and $\omega_t{}^{[sj]} \sim \mathcal{N}\left(0, R_t{}^{[sj]}\right)$ is the observation process noise. The observation process noise is same as the TOA estimation error $\varepsilon_{sj,t}$ defined in (2.17). It is assumed to be Gaussian additive noise.

The single target likelihood function for each transmitter receiver pair is then defined as:

$$p\left(\mathbf{z}_t{}^{[sj]}|\mathbf{x}_t\right) = \frac{1}{\sqrt{|R_t{}^{[sj]}|(2\pi)^l}} \exp\left(-\frac{1}{2}\left(\mathbf{z}_t{}^{[sj]} - h_{s,j}(\mathbf{x}_t)\right)' R_t{}^{[sj]^{-1}}\left(\mathbf{z}_t{}^{[sj]} - h_{s,j}(\mathbf{x}_t)\right)\right) \tag{4.56}$$

where for $h_{s,j}(\mathbf{x}_t)$ its first order linear approximation is used as in section 4.2.2. $R_t{}^{[sj]} = {\sigma_t^{[sj]}}^2$ is the measurement covariance matrix for the s, j transmitter receiver pair. l is the length of the observation $\mathbf{z}_t{}^{[sj]}$, which in range-only scenarios is 1, making $R_t{}^{[sj]}$ one-dimensional and thus just variance.

Combining the observations from all, $N_s \times N_{s,j}$, transmitter-receiver pairs, the likelihood function becomes:

$$p\left(\mathbf{z}_t|\mathbf{x}_t\right) = \prod_{i=1}^{N_s \times N_{s,j}} p\left(\mathbf{z}_t{}^{[i]}|\mathbf{x}_t\right). \tag{4.57}$$

The above likelihood function is applicable for single target scenarios. In the multiple target case, the case that a measurement originated from a specific target should be considered. That is, assuming independence of target states and measurements, the overall multiple target likelihood function can be defined as the product of the single-target likelihood functions.

Different single-target likelihood functions can be used as discussed in Section 3.4.2. However normal distributions have a wonderful property of closure under linear combinations i.e. the linear combination of two

independent normally distributed random variables also follows the normal distribution. This property is quite valuable for the derivation of the tracker equations. Thus if a differently defined likelihood function is used, then it can be approximated by GMs while keeping all other equations of the tracker same.

4.7 Method comparison and experimental results

The two approaches for multiple target multiple sensor localisation described above are verified on a scenario with two people moving perpendicular to each other with their paths crossing in the middle of the room. The person that has a shorter path turns and walks back. The same scenario and tracks are used in Section 2.6.2. A sensor network consisting of four UWB sensor modules is used. One of the sensors (Sensor 3 in Fig. 4.4) is placed behind one of the walls of the room, and the other three sensors are placed in three of the corners of the room. As mentioned earlier, each sensor consists of one transmitter and two receivers synchronised by an internal 7 GHz RF clock. About 25 impulse responses per second were measured. Directional horn antennas with different size and quality were used on all sensor nodes, resulting in varying sensor performance. No synchronisation between the different sensor nodes was considered, and all sensor nodes were running in parallel. The scenario is illustrated in Fig. 4.4, where the start position of the people is shown by a circle and the end position by an arrow showing the direction of movement the person had at the end of their movement.

The position of the transmitter and the two receivers of each sensor is shown in different colour for each sensor where the transmitter of each sensor is always placed between the two receivers of the sensor. The data used for target range estimation and consequently location estimation was obtained from a measurement campaign conducted at TU Ilmenau. The scenarios were measured indoors. Since GPS does not provide high accuracy indoors, the ground truth for these scenarios was extrapolated from the person motion path and positions at predefined anchor points at a given

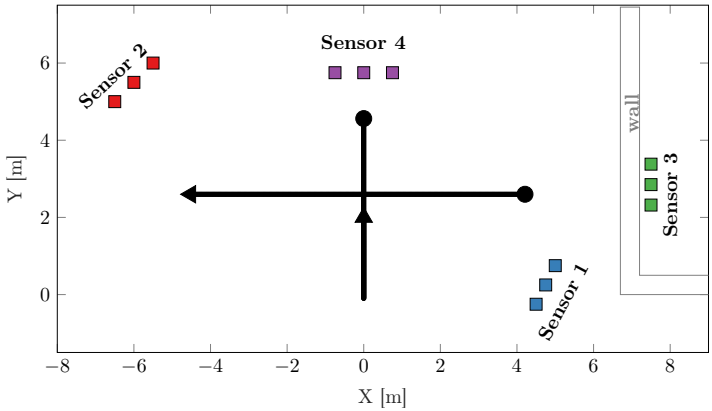


Figure 4.4: Measurement scenario used for method verification

time. Thus the ground truth used for comparison of the results is only a close approximation to the true ground truth path.

Fig. 2.15 shows the estimated locations of the two targets by each sensor. The different colours of the location estimates correspond to the sensors whose observations were used to determine the location. It can be noticed that the location estimates of each sensor are less noisy when the target is closer to the respective sensor. In Fig. 4.5 the fused location estimates are presented. As can be noticed, the fused location estimates are more accurate and less noisy than the direct location estimates per sensor and provide target tracks. The tracks of the direct localisation approach are shown in Fig. 4.6.

As can be seen, both two-step and direct approach seem to properly track both targets and determine their locations. There are minor visible differences between the results from the two approaches, which are quantified in Fig. 4.7. Here the optimal subpattern assignment (OSPA) [138] metric with cut-off $c = 10$ and order $p = 1$ is shown. The OSPA metric represents the sum of localisation and cardinality errors. It measures the distance

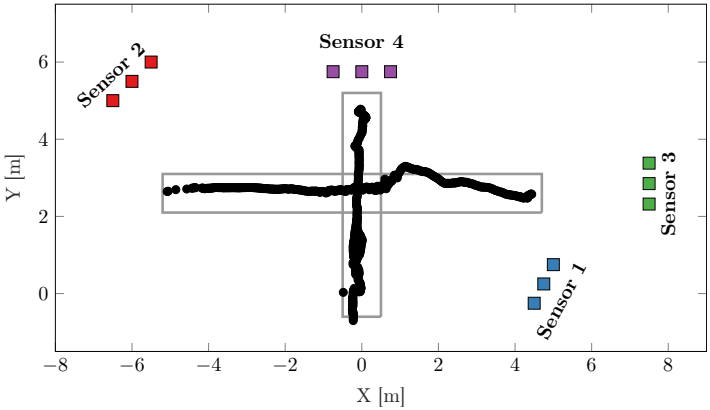


Figure 4.5: Location estimates fused by the fusion centre node

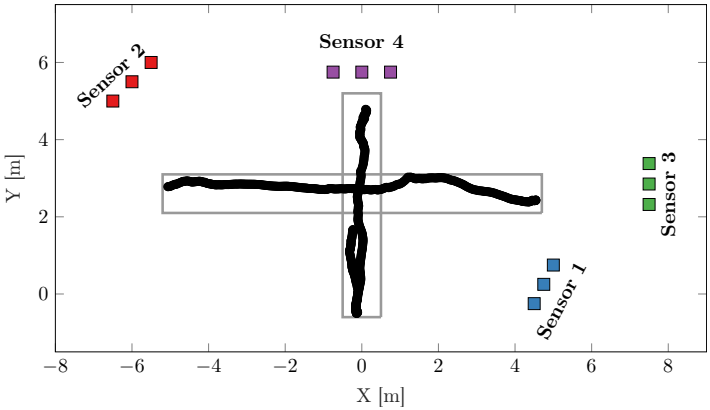


Figure 4.6: Target tracks using the direct range fusion approach

between two finite sets by minimizing the sum of distances, taking into consideration the set cardinalities. The cut-off parameter c deals with any mismatch between the number of elements of the two finite sets being

compared (the ground truth and estimated target states). A penalty c is added for each extra track in either set, representing either a missing existing target or an extra false target in the set of estimated targets. The maximum distance between two elements of the sets being compared should be less than c. Details on the OSPA metric are given in Appendix V.

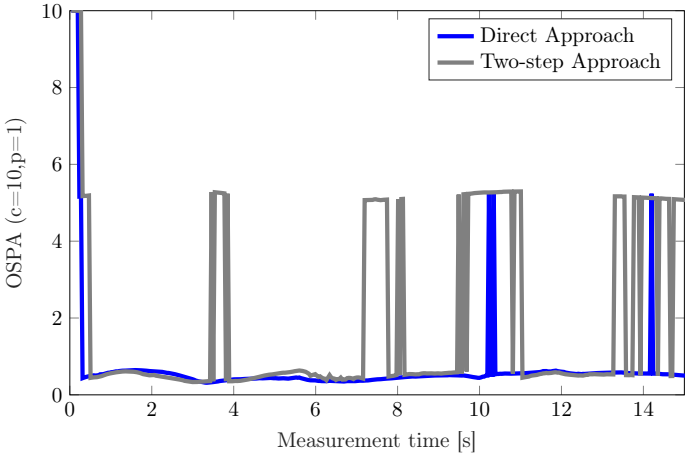


Figure 4.7: OSPA metric for the direct range fusion (blue) and two-step location fusion (grey) tracking approaches

Although both methods have rather low location estimation errors, the direct approach has greater stability. For both approaches, a measurement-based birth intensity was used, and all other parameters required by the GMPHD filter were kept the same, except for the differences in the type of observations and likelihood function as explaned in 4.6.2 and 4.6.3. This difference can be accounted to the fact that in case of target missdetection by one of the transmitter-receiver pairs of a sensor, no location estimate is provided for the fusion in the two-step localisation approach, whereas in the direct approach the existing range estimate of the other transmitter-receiver pair is used (see Fig. 3.9(d)).

4.8 Concluding remarks

This chapter introduces Bayesian target tracking methods. A brief overview and introduction to the different methods that may be used for single and multiple target tracking is given. In multiple target realistic scenarios complexities arise due to the uncorrelated observations available from both existing targets and clutter. The PHD filter is described as one possible solution to deal with these scenarios. The specific modifications for applicability of the PHD filter in multiple sensor scenarios is described using the sequential update equations in Section 4.5.2.

Two approaches for localisation of persons using bistatic range estimates by multiple UWB sensors are described. These approaches are simplistic, computationally solid and implemented for real-time operation. The first approach relies on single-sensor localisation being performed on the sensor node platform and fusion of the estimated locations using the GMPHD filter. In this case linear models are used. The second approach does only range estimation on the sensor platform, and a direct localisation and tracking GMPHD filter based fusion is performed. This means that the range observations of the targets are directly used as sensor observations in the multiple sensor PHD filter. Thus the observation model is a non-linear function. Linearisation as in the EKF equations is used.

The approaches are evaluated in near-real-time operation in a measurement scenario using UWB sensors. The measurement campaign was conducted in office environment which corresponds closely to real application scenarios for security or smart-home applications.

CHAPTER 5

OCCLUDED TARGET TRACKING

When used in realistic scenarios, UWB-based localisation becomes challenging due to errors caused by multipath propagation, multiple access interference, non-line-of-sight propagation, and others [139]. In most localisation and tracking techniques only positive information, i.e. target features observed by the sensors, is used. However, additional information is available from scenario areas and targets that should be observed by the sensor but no measurements are available due to possible occlusions of the targets present in these areas.

In this chapter localisation and tracking of multiple tag-free persons is considered as in the previous chapters of this thesis and presented in [1, 2], where persons are detected by the changes they impose in the IRF measured with a transmitter-receiver pair of a UWB module. In the presence of multiple persons, a person close to the transmitter or receiver of the sensor 'shadows' the persons located behind it with respect to the sensor since UWB signals are strongly attenuated after scattering from a person. Occlusions results in missing or incomplete measurements and are a serious challenge for extended multi-target tracking. An occlusion handling procedure can thus benefit the overall person localization and tracking procedure. This chapter presents such a handling procedure together with dynamic occlusion region modelling and incorporation within the tracking procedure presented in Chapter 4.

5.1 Introduction

Since a person is a complex target from radar point of view, UWB signals scattered from a person are strongly attenuated and often cause occlusion or missed detection of the other persons in the scenario. Shadowing influences on UWB sensors for multiple person detection and localisation are discussed in [140, 141] where weak signal enhancement, sensor positioning at higher altitude and use of multiple distributed sensors are suggested. Incorporation of the negative information in the multiple target tracking procedure is not considered.

Negative information describes the general case that no targets were detected in the field of view of the sensor where targets are expected. According to [142] measurements can be missing if an expected object is out of range, occluded or due to sensor failure. Thus if a target is within sensor range and there hasn't been any sensor failure, it can be concluded that the expected target is occluded either by other targets in the scenario or static objects. It is suggested that false interpretation of negative information can be avoided by modelling the measurement process as exactly as possible and considering occlusions by dynamic and/or static objects in the scenario.

Negative information is commonly used in robotics [143] in occupancy grid mapping. In [144], a centralised occupancy grid is generated from multiple cooperative sensors and used to confirm the tracking results by discarding tracks in unoccupied regions. [145] uses negative information to address sensor limitations such as Doppler blindness, jamming, finite resolution etc. In localisation, negative information is used in [142] for particle filter based localisation in a known map and in [146] for tracking known number of objects in a known map. Analytic model of occlusion is presented in [147], where a penalty cost is defined for unresolved tracks dependent on the modelled track visibility. In [148, 149] an occlusion likelihood model is derived for occlusions created by dynamic extended targets. A range-bearing sensor on an ego sensory vehicle is considered as source. Utilisation of the occlusion information improves the accuracy and

sensibility of state estimates for occluded objects. An early publication of the same authors, [150], was used as inspiration for the procedure described and applied within this work.

The PHD filter [53] provides good estimates of the target locations as long as a target is observed by a sensor. Track continuity and target death models have been proposed to handle missed detections, however they are mainly applicable when a target is not detected for very few consecutive time steps. When a target is not observed for multiple time steps, its weight falls below a predefined threshold and the target track is lost or discarded.

In [9, 10] dynamic target occlusion regions are introduced and incorporated within the target tracking system. Results show that fusing the occlusion information into a tracking filter improves the accuracy and sensibility of state estimates for occluded objects. In [9] a geometry based occlusion model was introduced and a simple occlusion handling procedure was applied. The results show that an occlusion handling procedure highly benefits the overall person localisation and tracking procedure. The models presented in [9] are extended in [10] where an occlusion likelihood function is derived and incorporated within a PHD filter based person tracking system. The proposed method is numerically and experimentally verified. This chapter reiterates and extends the work presented in [9, 10].

5.2 Occlusion region modelling

The occlusion region of a detected person is a region in the scenario where other persons/targets can not be detected due to the strong attenuation of the UWB signals caused by this person. For a known person location and extent and known sensor positions, the occlusion region can be modelled as the area behind the detected target with respect to the sensor as seen in Fig. 5.1.

Since only the range information of the target is used as observed by the UWB sensor, the location, extent and direction of that target need to be estimated and/or approximated from within the target tracking procedure.

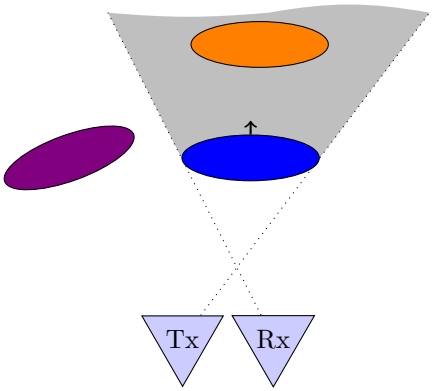


Figure 5.1: Occlusion region (grey) induced by the closest target (blue) - the orange target is shadowed and the violet target can be detected.

A person in 2D can be approximated as an elliptical target if the direction of movement is known.

In this section the occlusion region of a detected person i.e. the region of the scenario where due to the presence of this person, other persons can not be detected, is modelled. Since as target observation only the range is used, a unique solution for the target location based only on this one observation is not possible. If only range information is considered, a target n is occluded by a target m, if target m exists and the following range condition is satisfied:

$$R^{n,m} = (\tilde{r}_n > \tilde{r}_m) \tag{5.1}$$

where $\tilde{r}_m$ is the minimum range of object m with respect to a transmitter-receiver pair (red and green triangles in Fig. 5.2 and 5.3).

As can be seen in Fig. 5.2, based on this condition only the closest target can be detected and all others are within the occlusion region. However, some targets that satisfy Eq. (5.1) can actually be detected. Since bearing

observations are not available, the extent of a target needs to be defined or assumed and then used to limit the region that satisfies the above range condition. Thus, the target location is needed. The target location can be determined as the intersection of two ellipses (each defined by a target range estimate of different transmitter-receiver pair).

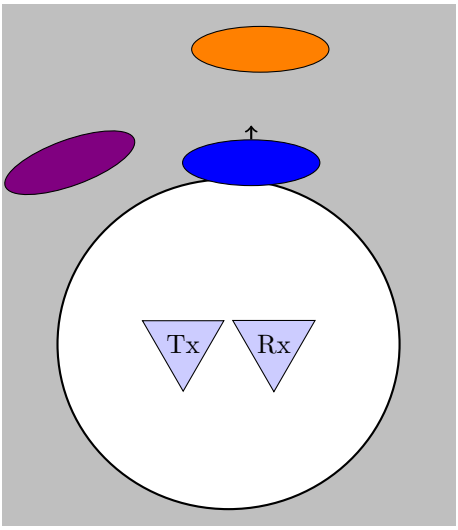


Figure 5.2: Occlusion region (grey) by using Eq. (5.1) - both violet and orange targets appear to be shadowed by the blue target if only the range condition is considered, although the violet target can actually be detected.

For simplicity, lets assume that a target m can be modelled by a circle with radius r_m and center (x_m, y_m) (blue dashed circle in Fig. 5.3). The range of the observed target is $\tilde{r}_m$ (represented by the black ellipse in Fig. 5.3). The target range is represented by an ellipse due to the bistatic configuration of the transmitting and receiving antennas. The target extent limits with respect to the sensor can be determined as the intersection points of the circle representing the target and the range ellipse (the two black points in Fig. 5.3).

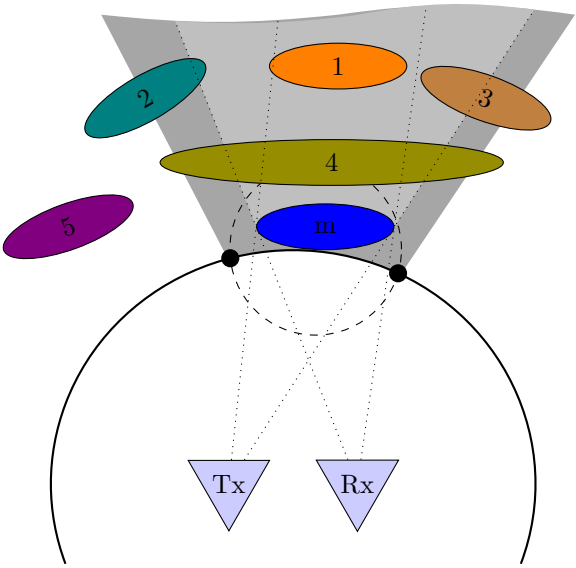


Figure 5.3: Occlusion region of a target: the modelled occlusion region (dark grey) based on modelled target extent and the true occlusion region (light grey) - target 1 (orange) is fully shadowed, targets 2, 3, and 4 are partially shadowed by target m (blue), and target 5 (violet) may be observed by the sensor

These intersection points can be determined by solving the two equations:

$$r_m = \sqrt{(x_m - x)^2 + (y_m - y)^2} \tag{5.2}$$

$$\tilde{r}_m = \sqrt{(x - x_s)^2 + (y - y_s)^2} + \sqrt{(x - x_j)^2 + (y - y_j)^2} \tag{5.3}$$

where (x_s, y_s) and (x_j, y_j) are the coordinates of the transmitter s and receiver j respectively.

To determine the region shadowed by target m, the left and right cut-off lines need to be determined. They are determined by the line passing through one of the antennas and one of the points determined as solution to the equations above (light grey dotted lines in Fig. 5.3). Two different regions can be determined. Often, one of these regions is not a convex polygon and can be thus discarded. If both regions are convex, the one that covers larger area is chosen as occlusion region. The modelled occlusion region is limited to the field of view of the sensor.

The true size, shape and orientation of the object influence the accuracy of this model. Any inaccuracies result in a conservative likelihood, guaranteeing that the true shadow is a subset of this approximation and is not mistaken to be visible.

In a generalised case, an arbitrary number of tracked targets exist, and each of them generates a shadow over the field of view. To avoid duplicates, the likelihood that target n is shadowed is defined as the likelihood that target n is shadowed by the target with minimum range to the sensor that also shadows target n.

5.3 Occlusion likelihood function

To determine the occlusion likelihood at a given point in time, the estimated location and extent of the present targets is needed. In the above section the occlusion region is modelled geometrically. In this section the minimum conditions that need to be satisfied to define the occlusion region analytically are given. This is similar to the modelled likelihood in [148], with the difference that in [148] both range and bearing measurements are available, and here only range measurements are available and the target extent is modelled to obtain its angular extent.

Object n is shadowed by object m if m is closer to the sensor compared to n, i.e. $R^{n,m} = (\tilde{r}_n \geq \tilde{r}_m)$, and the angular extend of object n is smaller than the extent of object m with respect to the sensor position, i.e. $B_1^{n,m} = (\theta_n \geq \theta_m) \cap (\phi_m \geq \phi_n)$, where θ_i is the clockwise angular extent

of object i and ϕ_i is the counter clockwise angular extent of object i. The angular extent of the objects is always calculated positive clockwise from the x-axis as shown in Fig. 5.4 where the target defining parameters are shown. The probability that object n is fully occluded by object m, $O^{n,m}$, assuming $R^{n,m}$ and $B_1^{n,m}$ are independent of each other, is then defined as:

$$p(O^{n,m}) = p(E^m)\, p(R^{n,m}|E^m)\, p(B_1^{n,m}|E^m) \tag{5.4}$$

where $p(E^m)$ is the probability that object m exists.

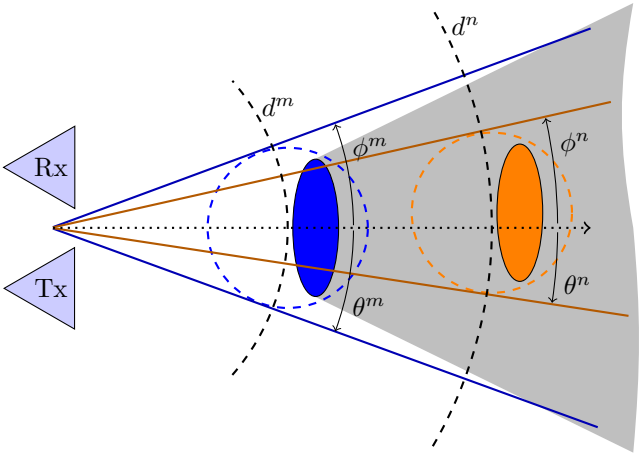


Figure 5.4: Full occlusion diagram of a target and defining parameters

The range and angular extent probabilities are computed using a smooth transition function based on the hyperbolic tangent function, i.e.

$$p(R^{n,m}|E^m) = s\left(\frac{\tilde{r}_n - \tilde{r}_m}{\sqrt{\sigma_{\tilde{r}_n}^2 + \sigma_{\tilde{r}_m}^2}}\right) \tag{5.5}$$

$$p(B_1^{n,m}|E^m) = s\left(\frac{\theta_n - \theta_m}{\sqrt{\sigma_{\theta_n}^2 + \sigma_{\theta_m}^2}}\right) s\left(\frac{\phi_m - \phi_n}{\sqrt{\sigma_{\phi_n}^2 + \sigma_{\phi_m}^2}}\right) \tag{5.6}$$

where

$$s(x) = \frac{1}{2} + \frac{1}{2}\tanh(kx) \tag{5.7}$$

with large k corresponding to a sharp transition at $x = 0$. Within this work $k = 1$ is used.

Different smoothing functions can be used such as Cauchy or normal distribution functions or any other cumulative distribution function of a continuous probability distribution peaked around zero and having a parameter to regulate the variance.

Analogous to the occlusion model in [148], the occlusion likelihood function is extended to also cover the cases of partially occluded targets such as targets 2, 3, and 4 in Fig. 5.3. For these cases additional angular extent conditions are defined i.e.

$$\begin{aligned} B_2^{n,m} &= (\phi_m \geq \theta_n) \cap (\phi_n \geq \phi_m) \\ B_3^{n,m} &= (\theta_m \geq \theta_n) \cap (\phi_n \geq \theta_m) \end{aligned} \tag{5.8}$$

where $B_2^{n,m}$ is the condition for partially occluded targets on the left of the occlusion region (covering case 2 and 4 from Fig 5.3) and $B_3^{n,m}$ is the condition for partially occluded targets on the right of the occlusion region (covering case 3 and 4 from Fig 5.3).

Since $B_2^{n,m}$ and $B_3^{n,m}$ both include case 4, they are not mutually exclusive. Thus the partial occlusion likelihood is

$$p(O^{n,m}) = p(E^m)\, p(R^{n,m}|E^m)\, p(B^{n,m}|E^m) \tag{5.9}$$

where

$$\begin{aligned} p(B^{n,m}|E^m) &= p(B_1^{n,m} \cup B_2^{n,m} \cup B_3^{n,m}|E^m) \\ &= p(B_1^{n,m}|E^m) + p(B_2^{n,m}|E^m) + ... \\ &\quad p(B_3^{n,m}|E^m) - p(B_2^{n,m} \cap B_3^{n,m}|E^m) \end{aligned} \tag{5.10}$$

and

$$\begin{aligned} p(B_2^{n,m}|E^m) &= s\left(\frac{\phi_m - \theta_n}{\sqrt{\sigma_{\phi_m}^2 + \sigma_{\theta_n}^2}}\right) \; s\left(\frac{\phi_n - \phi_m}{\sqrt{\sigma_{\phi_n}^2 + \sigma_{\phi_m}^2}}\right) \\ p(B_3^{n,m}|E^m) &= s\left(\frac{\theta_m - \theta_n}{\sqrt{\sigma_{\theta_m}^2 + \sigma_{\theta_n}^2}}\right) \; s\left(\frac{\phi_n - \theta_m}{\sqrt{\sigma_{\phi_n}^2 + \sigma_{\theta_m}^2}}\right) \\ p(B_2^{n,m} \cap B_3^{n,m}|E^m) &= s\left(\frac{\theta_m - \theta_n}{\sqrt{\sigma_{\theta_m}^2 + \sigma_{\theta_n}^2}}\right) \; s\left(\frac{\phi_n - \phi_m}{\sqrt{\sigma_{\phi_n}^2 + \sigma_{\phi_m}^2}}\right) \end{aligned} \tag{5.11}$$

In the general case of arbitrary number of tracked targets, N, the total occlusion likelihood of object n is

$$p(O^n) = \bigcup_{m=1, m\neq n}^{N} p(O^{n,m}) \tag{5.12}$$

Since an object n may be occluded by multiple targets simultaneously, the events in (5.12) are not mutually exclusive. Thus (5.12) can be expanded as in [148]:

$$p(O^n) = \sum_{i=1}^{N-1} \left\{ (-1)^{(i+1)} \sum_{j=1}^{{}^{(N-1)}C_i} \left[\prod_{k=1}^{i} p\left(O^{n,\mathbb{C}_j^i(k)}\right) \right] \right\} \tag{5.13}$$

where $\mathbb{C}_j^i(k)$ is the k-th object in the j-th i-combination and ${}^{(N-1)}C_i$ is the

number of i-combinations in the set of $N-1$ tracked targets (all targets excluding n).

The complement of the occlusion likelihood is the detection likelihood, i.e. the likelihood that target n is fully visible to the sensor

$$p(\overline{O^n}) = 1 - p(O^n) \tag{5.14}$$

If an existing track/target becomes unresolved over time, its estimate has large covariance. This uncertainty can be reduced by fusing the negative measurements and thus asserting that an unresolved target is unlikely to exist in areas that should be visible to the sensor.

5.4 Occlusion information integration

The occlusion model and likelihood function presented above are integrated within the PHD filter as shown in Fig. 5.5.

The **Feature extraction** block estimates the ranges from the raw measurements which are then used as observations in the tracker. The range estimation procedure is explained in Section 2.5. An observation dependent birth model for the target birth intensity is used as explained in Section 4.5.1. Each transmitter-receiver pair is considered as separate sensor and thus the multiple sensor update equations are used in the PHD filter. This is the way the conventional PHD filter is defined and used in Section 4.6.3. In the modified PHD filter the **Occlusion likelihood definition** block is used to define the probability of target occlusion based on the predicted target states. It is later integrated within the **Measurement update** block.

Often the probability of detection, $p_{Dt}(x)$ at time t is considered constant or dependent on the signal-to-noise ratio. In the modified version of the PHD filter a probability of target detection based on the occlusion likelihood is used. Thus, the detection likelihood is calculated as the complement of the occlusion likelihood defined in Section 5.3. The new probability of

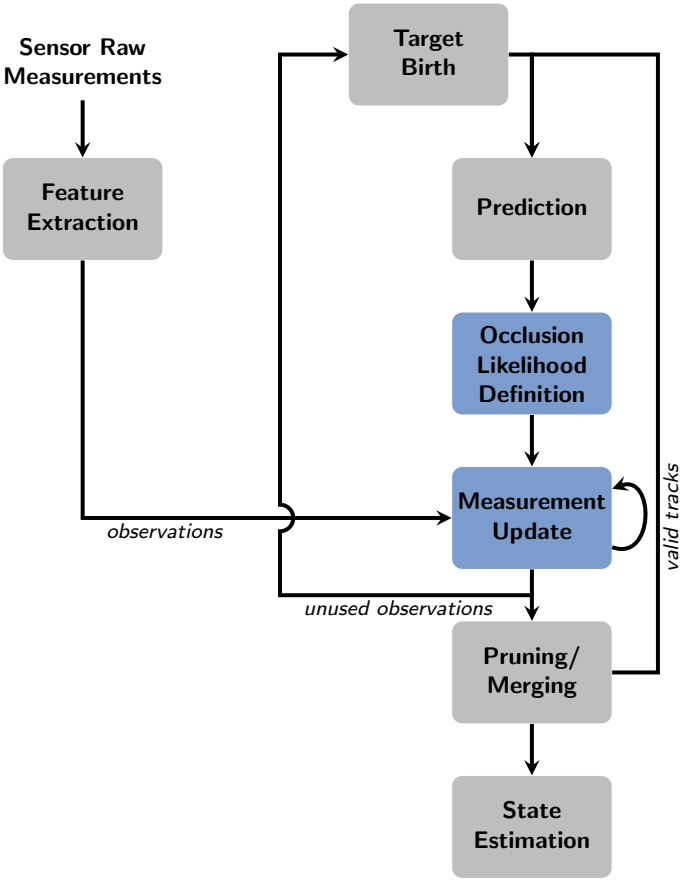


Figure 5.5: Block diagram of the modified multi-target tracker

detection is then defined as

$$p_{Dt}^{new}(\mathbf{x}_t) = p_{Dt}(\mathbf{x}_t)\ p(\overline{O^n}) = p_{Dt}(\mathbf{x}_t)\ (1 - p(O^n)) \tag{5.15}$$

This probability of detection is used in the update step of the PHD filter, i.e. equations (4.19), (4.46) and (4.47).

5.5 Evaluation and results

To verify the importance of negative information fusion and improvements it brings in multiple target tracking applications, the methods presented in this chapter are first applied on single sensor scenarios. First a simulation is used to verify the method applicability, and then the method is applied on a through-wall classroom scenario.

5.5.1 Numerical analysis

For demonstrating the performance of the described PHD filter with occlusion likelihood incorporation a scenario with three moving targets is simulated. The target trajectories are shown in Fig. 5.6 where the end state of each trajectory is indicated by a black circle. The sensor transmitter and two receivers are indicated by the black, blue and green triangles, respectively. For simulating the scenario a person is modelled as a circle with radius 0.3 m. Based on the sensor position and the target trajectories, the closest target shadows the other two targets at some points of time in the scenario.

The range observations of the targets including clutter is shown in Fig. 5.7, where the blue circles represent the observations with respect to the first receiver and the green squares with respect to the second receiver. Target 2 is shadowed by the Target 1 from time 29 to time 50. Target 3 is shadowed by the Target 1 and 2 from time 44 to time 65. $P(1)$ clutter points uniformly distributed along the range are simulated at each time step, where $P(x) = \frac{e^{-\lambda}\lambda^x}{x!}$ is the Poisson distribution.

The measurement noise standard deviation used is $\sigma_w = 0.1$ m. The probability of target detection by the receivers is assumed constant and same for both receivers, i.e. $p_{D_t}^{[j]}(\mathbf{x}_t) = p_D = 0.95$. Thus the probability of detection in the modified version of the tracker is $p_{D_t}^{new}(\mathbf{x}_t) = p_D\ (1 - p(O^n))$. For

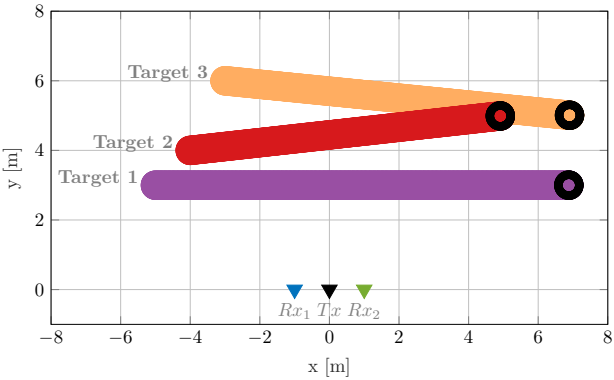


Figure 5.6: Simulated tracks for three targets (purple, red and orange) moving in a straight line ending at the respective black circle. Sensor position - transmitter (black) placed between two receivers, receiver 1 (blue) and receiver 2 (green)

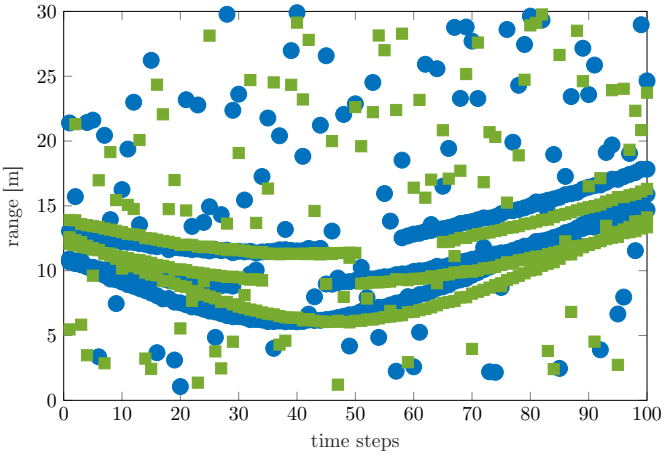


Figure 5.7: Targets observations (range as transmitter-target-receiver distance) by Rx_1 (blue circles) and Rx_2 (green squares)

the probability of occlusion, a target is modelled as a circle with radius 0.5 m. In both versions of the filter a process noise standard deviation $\sigma_v = 10^{-3}$ and probability of survival $p_{S,t}(\mathbf{x}_{t-1}) = p_S = 0.99$ is used.

In the target birth model the intersection points of all range-induced ellipses are used as x and y terms of the mean $m_{\gamma,t}$ of the newborn GMs. The $\dot{x}$ and $\dot{y}$ terms of the mean are 0. The covariance of the newborn targets is $P_{\gamma,t} = diag([1, 1, 0.1, 0.1])$. Pruning threshold of 10^{-4} and merging threshold 50 is used. The state extraction threshold is set to 0.5.

To evaluate the performance of the PHD filter the OSPA metric with cut-off $c = 10$ and order $p = 1$ is used [1]. The modified PHD filter with occlusion likelihood incorporation and the conventional sequential PHD are compared in Fig. 5.8 where the OSPA metric averaged over 100 Monte Carlo runs for the scenario described above is shown.

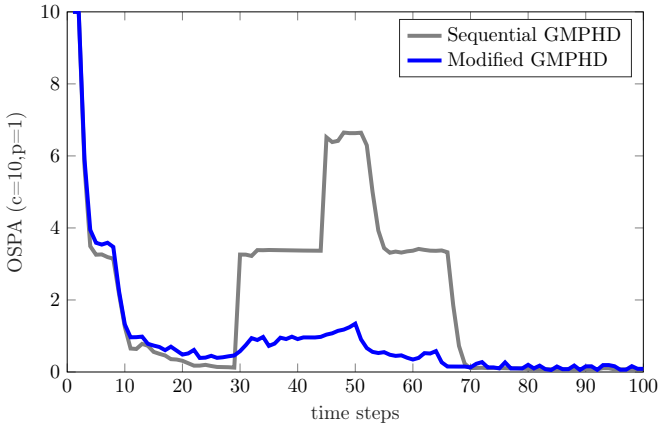


Figure 5.8: Average OSPA metric using the sequential PHD filter (grey) and the modified PHD filter with occlusion likelihood incorporation (blue)

As can be observed, using the PHD filter with occlusion likelihood to

[1] Details in Appendix V and [138]

define the probability of target detection leads to better results, since the filter is able to track the occluded targets (between time step 29 and 65 one or two targets are occluded) and update their state as soon as there is a fitting observation of the previously occluded target. It can be observed that from time step 65 to 69 although all three targets are detected, the GMPHD takes some time until it can correctly estimate the states of all three targets, whereas the GMPHD filter with occlusion incorporation is quicker in correcting the track.

5.5.2 Experimental verification

The suggested modification with occlusion likelihood integration is applied to data acquired using one UWB sensor node with transmitter and two receivers placed behind a wall. The scenario is as described in Fig. 5.9. Two persons walk in a room along roughly predefined paths straight to and from the sensor node to the end of the room. The complete ground truth path of the motion of the persons is unfortunately not available and thus a close approximation is used. The UWB module used has 3.5 GHz bandwidth. The measurement rate is 100 IR/s.

The feature extraction procedure is range estimation from the received IRFs at each time step as in Section 2.5. The background subtraction method is exponential averaging with forgetting factor 0.85 which removes the static background reflections. For range detection the CFAR detector with false alarm probability 0.15 is applied due to the low signal-to-noise ratio resulting from the signal attenuation when propagating through the wall. To reduce the number of range estimates per target and transmitter-receiver pair to one a hierarchical clustering algorithm is applied. Range tracking is not used at this time. The estimated ranges from the IRF received by receiver 1 are shown in Fig. 5.10. The expected range estimates at each time point from a simulation of the same scenario are depicted as grey line for comparison. As expected, the person closer to the sensor is detected whereas the other person is only sometimes detected. This is accounted to the low signal-to-noise ratio, the attenuation of the UWB signal

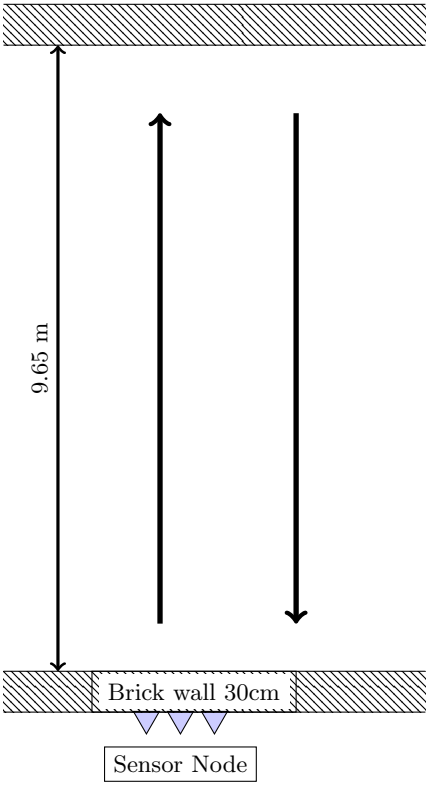


Figure 5.9: Measurement scenario schematics

due to the penetration through wall, and in part to the shadowing imposed by the person closer to the sensor over the rest of the scenario. Based on the approximated ground truth the person closer to the sensor partially occludes the other person throughout the complete measurement time. However, in the first 3 s of the track less than 50 % of the body is occluded, and thus it can often be detected by the sensor. Starting from the 4^{th} second larger portions of the body (more than 50 %) are occluded increasingly.

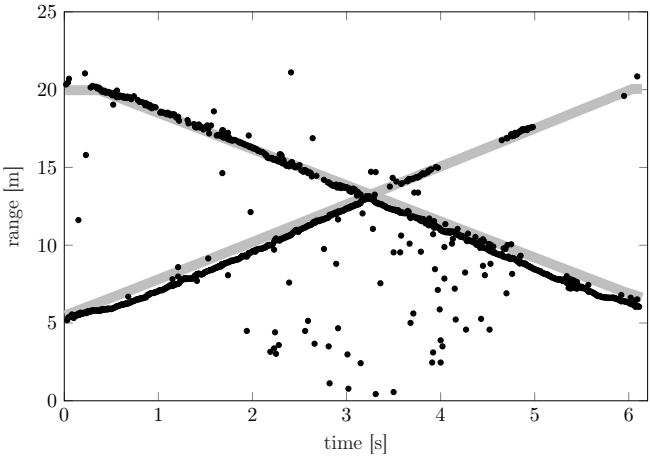


Figure 5.10: Target range estimates using CFAR and simulation based expected range estimates (grey line) for receiver 1 in the two person scenario

Detected targets are modelled as circles with radius 0.7 m and the angular extent parameters are obtained from the model at each time point for each detected target. The radius used to model the target is very large, however it partially accounts for location estimation errors of the detected target. In this scenario location estimation errors arise due to the vicinity of the antennas. The range estimates of both transmitter-receiver pairs are close to each other and the intersection of the range induced ellipses is flat. The accuracy of the occlusion likelihood estimation is highly dependent on the location estimation of the detected targets. Location estimation errors of detected targets can result in predicting high occlusion probability for non-occluded targets or predicting low occlusion probability for occluded targets. Modelling the target with larger radius or based on the error covariance of the current estimate might lead to better results and is worth investigating in future. It should be noted that since range-only observations are often not enough to accurately localise the targets, the error covariance can be rather

large resulting in targets appearing to shadow the whole scenario. Another option would be to model the targets as ellipses by using the estimated target heading.

For this scenario measurement noise standard deviation $\sigma_w = 0.3\,\mathrm{m}$ and process noise standard deviation $\sigma_v = 0.5$ is used. The other filter parameters are same as in the simulation scenario.

It should be mentioned that in the implementation of both conventional and modified GMPHD filter, the number of estimated targets equals the number of surviving mixtures after merging and pruning, and the state estimate of these targets is the mean of those mixtures. Thus, as expected when the paths of the two targets are crossing (around 3.5 s), the filter doesn't manage to keep the tracks of both targets.

The OSPA metric with cut-off $c = 10$ and order $p = 1$ of the scenario for the sequential GMPHD filter using only target range estimates and the modified GMPHD filter using the modified probability of detection based on the occlusion likelihood in addition to the same range estimates is shown in Fig. 5.11 [2].

As can be seen in Fig. 5.11, the OSPA metric of the conventional GMPHD filter results in many jumps over the measurement time. This can be attributed to the many target missed detections due to it being partially occluded by the other target. In the first half of the measurement time a smaller portion of the target is occluded and thus it is often detected. However there are still many intervals where the target is not detected (missed detections). Since the PHD filter is not a conventional tracker, if a target state is not updated by an observation for multiple time steps it's weight falls below the estimation threshold and it's state is thus not part of the estimated target states.

Sample of the scenario with the occlusion likelihood at the given time step shown in the background and the estimated target locations as green error ellipses for the conventional and modified GMPHD filter are shown

[2] Details on the OSPA metric can be found in Appendix V and in [138]

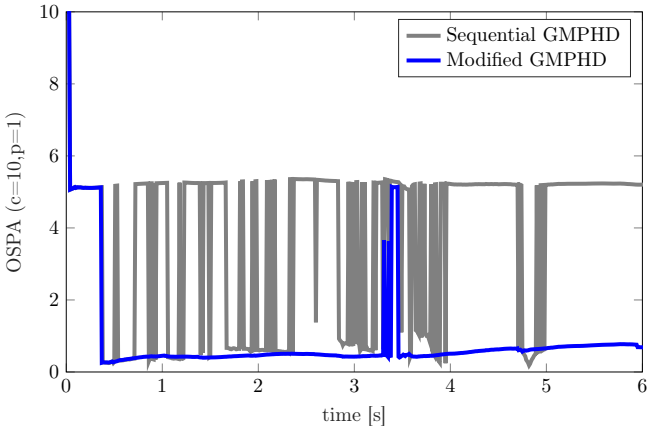


Figure 5.11: Average OSPA metric using the sequential GMPHD filter (grey line) and the modified GMPHD version with occlusion likelihood incorporation (blue line)

in Fig. 5.12. The black circles represent the approximated ground truth location of the persons. As can be seen in this sample, integrating the occlusion likelihood within the PHD filter allows for tracking both targets in the scenario although one of them is occluded by the other and does not provide observations to the sensor. In the left figure, the target locations are estimated using the sequential PHD filter without the use of the occlusion likelihood. The occlusion likelihood is depicted there for comparison.

As shown in Chapter 3 using multiple UWB sensor nodes improves the localisation and tracking of multiple persons. However in this chapter it is shown that by using a single sensor and considering the occlusion likelihood, multiple persons may be localised and tracked. Having multiple sensors observing the scenario is obviously an advantage, however, there might be scenarios where placing multiple sensors might be impractical or restricted. In such cases without the incorporation of the occlusion likelihood, multiple targets can not be tracked.

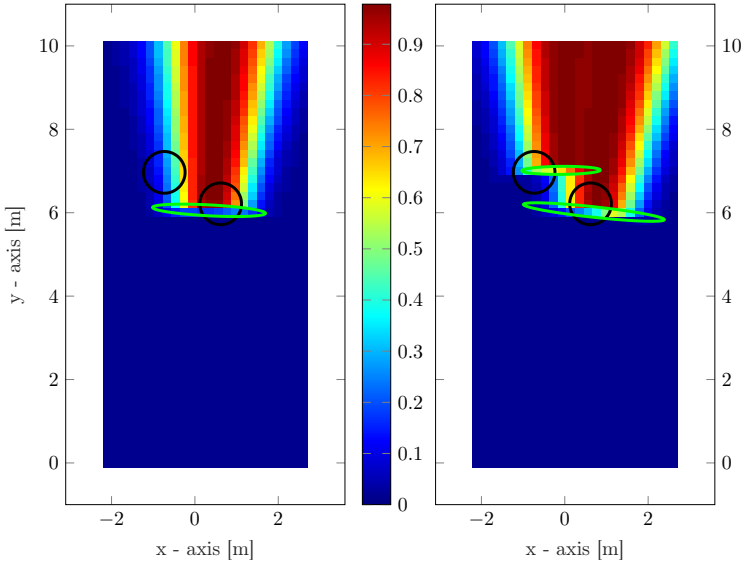


Figure 5.12: Sample occlusion likelihood at a given time point calculated for the scenario without (left) and with (right) occlusion likelihood incorporation. Black circles represent the targets positions and green ellipses representing the estimated state covariances of detected targets.

5.6 Concluding remarks

In multiple person localisation scenarios, person-induced shadows (or occlusions) lead to target missdetection. In this chapter an occlusion model for dynamically defining the occluded area in the scenario at each time point based on the estimated target locations is presented. Based on this model, an occlusion likelihood function is described and is later used to calculate the probability of target detection. The modified GMPHD filter is evaluated on a simulated scenario with three moving persons showing a significant improvement compared to the typical GMPHD filter where occluded targets are simply discarded after few time steps of non-detection.

The modified GMPHD filter is also applied on experimental data gathered using a bat-type UWB sensor. The improvement in target localisation with the modified GMPHD filter is significant since occluded targets continue to be tracked even tough there are no detections from the sensor. Future improvements by incorporating the occlusion likelihood function in the innovation would limit the error estimate within the occluded region. In addition improvement of the target extent model should result in improved target estimates with lower error covariance.

CHAPTER 6

CONCLUSION AND FUTURE WORK

Localisation and tracking of people using UWB sensors is important in many applications for indoor localisation. This thesis concentrates on defining a real-time framework for localisation and tracking of multiple people in realistic scenarios. On one side the use of multiple distributed sensors is explored and on the other side potential improvement due to negative information fusion is investigated.

6.1 Summary

Indoor localisation of people that do not cooperate with the system network can be achieved and complemented by the use of UWB systems. Due to their high time resolution, various multipath components can be separated from the echoes of the people of interest. The large frequency spread makes UWB systems applicable for through-wall detection and localisation by using the lower frequencies of the spectrum.

This thesis focuses on detection, localisation and tracking of people using UWB sensors for real-time applications. First and foremost the basis of UWB sensing is described, followed by the operating principle of single UWB sensor for person detection and localisation. Different methods are discussed by concentrating on most fitting methods for multiple person detection and localisation in real-time for unknown scenarios. Since a person is an extended

target for UWB sensors in indoor scenarios, multiple detections per person should be considered. Within this thesis a hierarchical clustering approach is used over the target detections for reducing the multiple observations per sensor per target to one. A range association technique is proposed to associate estimated ranges to existing targets for simple single sensor based localisation.

Once the ranges are associated, single target localisation can be used. Since each target range corresponds to an ellipse in Cartesian coordinates, the target location is derived as the intersection of the ellipses derived by associated range estimates. This method has its drawbacks in the scenarios and sensor constellations considered within this thesis. Since the two receivers of a sensor are close to each other the ellipse intersection (or annuli intersection if range estimate errors are considered) is rather flat and leads to large location estimation error. Additionally, if only one range is associated to a target it is discarded and can not be used for target localisation.

In multiple person localization scenarios, person-induced shadows lead to target missdetection. Thus the question of how to better detect, localise and continuously track multiple people in the area of interest is raised. Two directions are explored in this thesis.

The first direction is in expanding the field of view of the system by using multiple distributed sensors. The concept, structure and architecture of a distributed UWB sensor network is described. In addition to allowing for detection of possibly shadowed targets in the area with respect to one sensor, it also increases the coverage area of the system and can lead to more accurate detection and localisation of multiple people in the overlapping areas. The processing of the measured impulse responses is also distributed in the network for reduction of the data transmission load between the network components. Based on the concept developed within this thesis a simple demonstrator is implemented.

Fusing information from multiple sensors can be tricky. Typical multiple

sensor fusion challenges are described and a multiple sensor likelihood function for range-only detections is derived.

For improving target detectability over time, target propagation is considered. Brief overview of Bayesian tracking and some derived methods for single and multiple target tracking is given. More detailed explanation of the PHD filter as a filter of choice is provided with additional explanations regarding the target birth model, multiple sensor application and observation models used for successful application of the filter in real time over the information obtained by our sensor system.

This thesis describes two approaches for localisation of people using the measured target scatterings by the multiple sensor receivers in the sensor network. The described approaches are simplistic, computationally solid and implemented for real-time operation. The first approach relies on single-sensor localisation being performed on the sensor node platform and fusion and tracking of the estimated locations at the fusion centre node. The second approach does only range estimation on the sensor platform, and a direct localisation and tracking is performed on the fusion centre platform. An extensive multiple sensor tracking framework based on the PHD filter is proposed with measurement-based target birth model and sequential measurement update procedure. Both methods have low location estimation errors, however the second approach has greater stability over time.

The second direction explored is to use the missdetections of targets by a sensor to our advantage by analysing the reasons of missdetection and if the cause is due to shadowing of the targets of interest either by each other or by other objects in the scenario, to incorporate this information into the multiple target tracker. A dynamic occlusion model is derived based on an assumed target extent, followed by derivation of an occlusion likelihood function. This occlusion likelihood function is used to redefine the target probability of detection dynamically before it is used in the tracking framework. The improvement in target location estimation when using

negative information is significant since occluded targets which can not be detected by the sensor continue to be tracked.

The approaches are evaluated in near-real-time operation in measurement scenarios using M-sequence UWB sensors, conducted in office scenarios which correspond closely to real application scenarios for security or smart-home applications.

6.2 Future work and extensions

During the course of this thesis different research questions arose which due to time constraints could not be covered within this work and are thus proposed as future directions of research.

The background subtraction method used here for suppression of stationary reflections is very useful for online real-time applications. However contributions of stationary people are also suppressed. Using UWB sensors, heartbeat and breathing patterns of people can be detected [151, 152]. Thus it would be preferable to develop a detection method by which both moving and stationary people can be detected.

The clutter present in the detections seems to often be following the movement of one of the targets. This may be delayed target reflections due to additional scattering from walls (multipath). Thus a filter for removing them may be developed, or they may be also included as additional target observations. The second suggestion has been explored in [13] for single person detection and localisation (behind a corner scenario). Multipath propagation exploitation can also provide additional information when the number of sensors observing the targets is small. A ray-tracing engine can be used for multipath prediction.

The sensor network system and approach can be further improved by inclusion of other parameters corresponding to moving people, which can be extracted from the measured impulse responses. Hardware constraints limited the explorations of this topic. TDOA and Doppler information can

greatly improve the localization and target characterization/classification approach. In addition a higher cooperation level between the sensors can lead to additional information resulting from using signals transmitted by one sensor node and received by another sensor node as described in the concept presented in Section 3.2. Different sensor placement geometries can be explored for finding an optimal sensor geometry for different scenarios.

The tracking framework presented here can be expanded for extended target tracking and thus observation inaccuracies due to possibly over simplistic clustering can be avoided. The occlusion likelihood derived in Chapter 5 can benefit from an appropriate target extent modelling. Incorporating the occlusion likelihood fully in the update process of the tracker can limit target covariance estimates of occluded targets to occlusion regions only.

An interesting field to explore in is target classification using UWB sensors. Different polarisations as well as micro-Doppler information of the detected targets may be used for classification. Categorisation of the moving objects in the defined scenario as human or non-human can help in simplifying the localization and tracking procedures by considering only targets classified as human in those procedures. This would be helpful in rescue scenarios where detection and localisation of people is a priority. A simple and effective solution to this problem can be based on waveform analysis where certain statistical features can be extracted from the return signals, and then compared to a database of feature sets from different objects. This procedure requires statistical models of different objects that can be encountered in the scenario for which a lot of training data is required. It also requires training data from different people to cover the diversity of human forms.

Appendices

APPENDIX I

LOCALISATION SCHEMES AND SYSTEMS

Different classification schemes for localization technologies exist. One important classification is based on the signalling scheme it uses. RF frequency is one of the most commonly used signalling scheme for localization purposes due to its ability to penetrate through obstacles and propagate long distances. Infra-red signals have low power and are inexpensive but can not propagate through obstacles and are susceptible to sunlight thus requiring an extensive network of infra-red sensors in the indoor environment to pick up the transmitted signal. Optical signals also require LOS conditions, are affected by sunlight and require low power. High accuracy can be achieved and these sensors are appropriate for short ranges (less than 10 meters). Ultrasound signals provide high accuracy in short range. Signals travel slowly and thus slow clock is sufficient and high accuracy can be achieved cheaply in LOS conditions. However acoustic emitters need a lot of power and do not work well in NLOS conditions.

Depending on where the localization takes place, there exist handset-based (location aware) and network-based (location support) systems. In the first one the target calculates its own position based on the signals it receives from the reference nodes, also known as self-positioning. In the later one a central processing unit calculates the targets position, also known as

remote positioning. Privacy issues may be of concern for the later one.

Localisation systems can also be classified as indoor or outdoor systems due to the significant propagation characteristic differences in each environment. Examples of outdoor localisation systems are: GPS using more that 24 satellites with location accuracy of 1 to 5 meters, or the E911 service in cellular networks. Both perform well outdoors but are not designed for indoors where technical challenges exist and accuracy requirements are typically much higher. Indoor systems may require unique infrastructure or use pre-existing infrastructure such as wireless local area network (WLAN). Software-based localization can be implemented by using existing infrastructure and there is no need to deploy additional hardware for localization purposes. Hardware based localization systems require installation of additional hardware.

In active localization systems the network sends specific signals for location estimation of the targets. In passive localization systems pre-existing signals are used, i.e. no specific signal for localization purposes is transmitted.

In centralized localization systems, position related information is forwarded to a data fusion center where the target's location is estimated. Terminals that use decentralised localization systems determine their location jointly by communicating with each other.

Additionally, if the target to be localised cooperates with the localisation system, for example carries an RF tag, the estimation process is also termed cooperative localisation. Otherwise if the target is just a passive bystander and is possible not aware it is being localised, or it can not help the localisation system, the estimation process is termed non-cooperative localisation.

Accuracy is the metric that tells us how well a certain localization technology performs. It is defined as how far the estimated location of a target is away from its actual location. It is desired that a certain localization accuracy is achieved with high probability. Precision defines the percentage that a certain accuracy or better is achieved. In cell identification

localization, the target's position is approximated based on the location of the reference node, limiting the accuracy to the cell size. In triangulation based systems, the intersection of at least two lines obtained from AOA information from at least two reference nodes are used. In trilateration systems at least three reference nodes are needed for 2D localization and at least four nodes are needed for three-dimensional (3D) localization. Circle or hyperbola intersections obtained from TOA, RSS or TDOA are used for location estimation. Fingerprint-based or pattern-matching localization technologies compare real-time measurements with a location database to infer the terminal location. Accuracy is limited to the granularity of the training locations and an off-line calibration stage is required which may need to be repeated if the environment propagation characteristics change. Range-based systems depend on the distances (TOA, TDOA, RSS) or angles (AOA) between the nodes. Range-free localization systems do not require estimation of the absolute distances. They either rely on high density of anchors (i.e. location estimate is the average location of all connected anchors) or are hop counting techniques.

Within this thesis indoor localisation of non-cooperative targets using UWB radar is considered. This same problem description is often wrongly referred to as passive localisation. A two-step trilateration-based localisation system is to be used, where in the first step target's ranges are estimated and in the second step they are used for target localisation.

APPENDIX II

PARAMETER ESTIMATION ACCURACY

In two-step localisation techniques, the first step is to estimate signal parameters such as RSS, AOA, TOA and TDOA, based on which the target location can be estimated using triangulation, trilateration or hybrid techniques. Typically single parameter can be estimated for each received signal, however if feasible multiple parameters may also be estimated to improve accuracy. Below, the estimation accuracy of some of these parameters is discussed.

Received signal strength. RSS measurements provide distance information between two nodes by measuring the power of the received signal. The average signal power decays with distance. UWB signals experience multipath fading, shadowing and path loss (PL) while propagating. Ideally signal power over longer time interval is not as influenced by the multipath fading and shadowing, and thus the model is as follows:

$$\overline{P}(d) = P_0 - 10n \log_{10}(d/d_0) \tag{II.1}$$

for PL exponent n, distance d between nodes and P_0 average received signal power in dB at reference distance d_0. The PL exponent is hard to estimate due to propagation influences such as reflection, scattering and

diffraction. The multipath effect in UWB systems can be mitigated by having long enough integration interval to include all multipath components. In the case of tag-free target detection, in addition to multipath fading, shadowing and PL, backscattering of the object and its RCS should also be considered. If the received power $P(d)$ is modelled as a Gaussian random variable with mean $\overline{P}(d)$ and variance σ_{sh}, the CRLB for range estimation using RSS can be expressed as

$$\sqrt{Var(\hat{d})} \geq \frac{\ln 10\sigma_{sh}}{10n} d \tag{II.2}$$

with $\hat{d}$ being the unbiased estimate of d and σ_{sh} the standard deviation representing the log-normal shadowing effect. For transmitter detection, the standard deviation of the error is always above $1m$ for distances above $6m$, meaning that RSS measurements do not provide very accurate range estimates for UWB systems [65].

Angle of arrival. AOA is the angle between the propagation directions of the incident waveforms. For measuring the AOA antenna arrays are used, and the angle information is obtained by measuring the difference in arrival time of the incoming signal at different antenna elements. The idea is that the difference of the arrival times at different antenna elements contains the phase information for a known geometry. For narrowband signals time-difference can be represented as a phase shift and thus the angle can be estimated by testing phase-shifted versions of the received signal at the array elements. Due to the large bandwidth in UWB, the time delay can not be represented as a unique phase value and thus time-delayed versions of the received signal should be considered.

Time of arrival. The TOA parameter measures the time needed from the target node to the reference node in active localisation and the time needed for a transmitted signal to reach a receiver by bouncing (scattering) off a target in passive localisation. To estimate unambiguous TOA parameters the transmitter and receiver need to have a common clock or

exchange time information. In bistatic configuration for passive localisation techniques, each TOA measurement forms an ellipse. For unambiguous 2-D positioning, at least three synchronised transmitter-receiver pairs are needed.

The conventional TOA estimation technique uses matched filtering or correlation to estimate the time shift of a transmitted signal. For a simplified received signal model

$$r(t) = \alpha s(t - \tau) + n(t) \tag{II.3}$$

where τ represents the TOA, α represents the channel coefficient, $n(t)$ is white Gaussian noise with zero mean and $\mathcal{N}_0/2$ spectral density, and $s(t)$ represents the transmitted signal, the CRLB of the TOA measurement error for single-path AWGN channel can be expressed as:

$$\sqrt{Var(\hat{\tau})} \geq \frac{1}{2\sqrt{2}\pi\beta\sqrt{SNR}} \tag{II.4}$$

where $\hat{\tau}$ is the unbiased TOA estimate and β is the effective signal bandwidth:

$$\beta = \left(\frac{1}{E} \int_{-\inf}^{\inf} f^2 |S(f)|^2 \right)^{1/2} \tag{II.5}$$

with $S(f)$ being the Fourier transform of the transmitted signal and E the energy of the signal. The SNR is $SNR = \alpha^2 E/\mathcal{N}_0$.

As can be noticed, TOA estimates can be improved by increasing the SNR or the effective bandwidth β whereas RSS cannot. Since UWB signals have large bandwidth, very accurate TOA parameters can be estimated.

Time difference of arrival. If clock synchronisation between the transmitter and receiver is not available, TDOA parameters may be estimated if multiple receivers are synchronised. TDOA represents the difference in the signal arrival time between two synchronised receivers. TDOA estimates can be obtained by estimating the TOA at each receiver (with

time-offset) and calculating the difference between the two TOA estimates. Analogous to TOA, TDOA estimates may be improved by increasing the SNR or signal bandwidth. TDOA may also be obtained by cross-correlating the received signals from the two synchronised receivers to calculate the delay corresponding to the largest cross-correlation value. This method wors well in single-path channels but its performance degrades in multipath channels and in the presence of coloured noise [65].

Doppler frequency. The Doppler effect is often considered to be pure frequency shift (also referred to as Doppler shift or Doppler frequency) caused due to motion of a target. This interpretation is however only valid for narrowband signals and does not hold for wideband signals [61]. When affected, signals are either stretched or compressed, which in narrowband case corresponds to shift of the center frequency of the signal. In wideband case the transmit and receive signals decorrelate with increasing speed and signal duration [61]. To determine the Doppler effect, the ambiguity function should be sensitive to both time shift and Doppler scaling. The width of an ambiguity function in time-delay or velocity can be controlled by the bandwidth and duration of the sounding signal. Narrowband systems tend to have bad range (low bandwidth) and good Doppler (long signal duration) resolution whereas wideband systems tend to have good range (large bandwidth) resolution and bad Doppler (short signals) resolution. In narrowband processing the decorrelation between the receive and transmit signal does not have to be respected. The Doppler ambiguity is mainly determined by the system update rate. For a system measuring $25IR/s$ the refresh rate of the system is $25Hz$ which corresponds to maximum detectable Doppler frequency of $\pm 12.5Hz$. However in wideband processing a two-dimensional search in time-delay and Doppler scaling is needed which increases the complexity of the sensor electronics. Typically one tries to avoid Doppler influence by either using signal which are not sensitive to Doppler (short pulses) or by using signals which do not decorrelate (e.g. linear frequency modulated sine waves). More information on UWB systems and designs can be found in [61] and on narrowband and wideband ambiguity functions in [153].

APPENDIX III

RESULTS COMPARISON IN A SAMPLE SCENARIO

By using a distributed sensor network for detection and localisation of targets in an indoor scenario more information is obtained which can be used for more reliable target localisation. This information abundance helps in consistently localising targets even in the event of target missdetection by one of the sensors. Although obvious, the information gain by using multiple sensors has been verified experimentally in realistic indoor scenarios.

Lets consider the scenario from Fig. 3.6 and the centralised fusion approach. The target range is estimated using the methods in Section 2.5. For the background subtraction method a forgetting factor of 0.85 is used. The false alarm probability for the CFAR detector is set to 0.1 due to the low SNR. Hierarchical clustering is applied on the CFAR estimates to reduce the number of range estimates per target to one (due to the target being an extended target for the system). The estimated ranges of the two receivers of the first sensor are fused together using the sequential PHD filter described in Section 4.6.3. The resulting location estimates (tracks) are shown in Fig. III.1a in red. The estimated ground truth tracks are shown in black. As we can see, by using only one sensor we can not track the full track of the person. By fusing all range estimates from both sensors, the targets can be tracked continuously over the full time span of the scenario.

The results of the fused sensor information can be seen in Fig. III.1b. The same thing happens when two persons are present in the area of interest. The target tracks when using a single sensor and when using both sensors are shown in Fig. III.2. The information gain by using multiple sensors is obvious. Both targets can be tracked throughout the whole scenario when using both sensors.

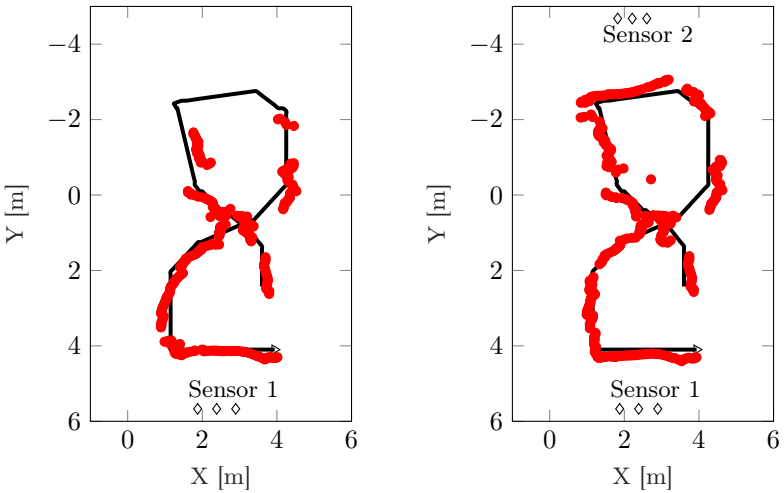


Figure III.1: Estimated target track (red) by using (a) single sensor and (b) fusion of the estimates by both sensors. The corresponding estimated ground truth track of the person is depicted in black.

The presented results show that by only using a single sensor there are cases when there is not enough information to localize multiple moving persons. When only one sensor is used the targets further away from the sensor are not detectable mainly due to them being shadowed by the target which is closer to the sensor. This same scenario is used in Section 5.5 where instead of using two sensors, negative information fusion is used to

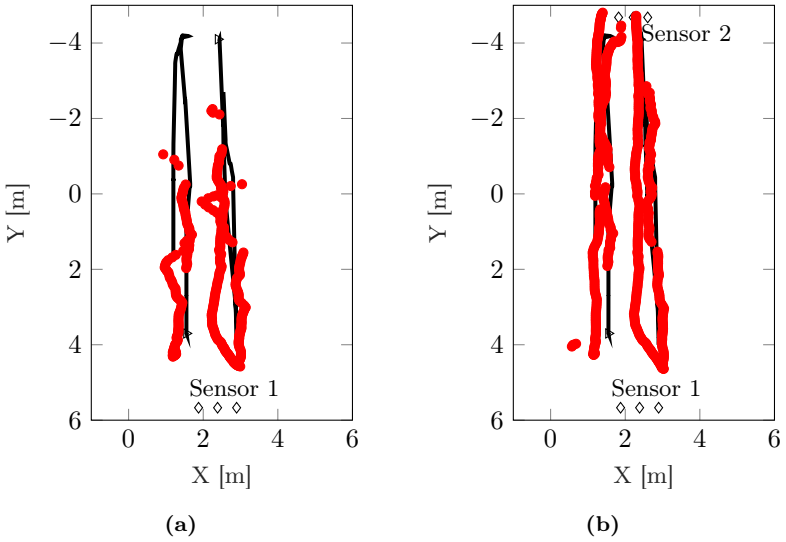


Figure III.2: Estimated target tracks (red) by (a) using only sensor 1 and (b) fusing information from both sensors. The corresponding estimated ground truth track of the persons is depicted in black.

improve the target location estimation.

APPENDIX IV

CENTRAL AND DISTRIBUTED DATA FUSION

The main idea as well as challenge of dynamic sensor fusion is the fusion of data from all the sensors (observing a phenomenon represented as a random variable $x(t)$) to provide an estimate.

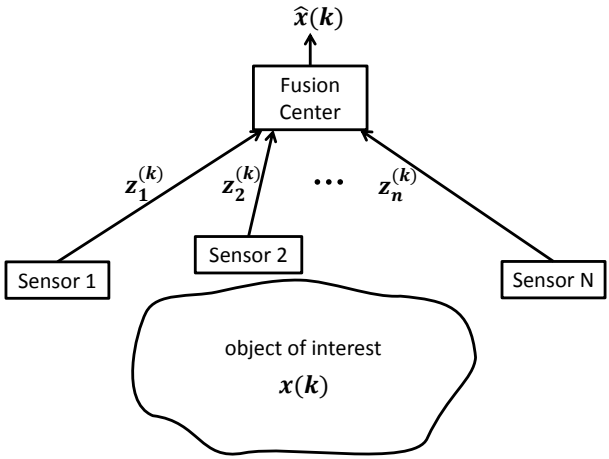


Figure IV.1: Distributed state estimation

If all sensors transmit their measurements at every time step, the problem can be solved. The theoretically optimal solution is the fusion of all available observations at a global fusion centre (centralised fusion). Alternatively distributed tracking may be used where tracks are locally formed at the nodes and these tracks are then fused at the fusion centre (distributed fusion). The problems of the centralised fusion is firstly the computational cost i.e. the central node needs to handle matrix operations that increase in size as the number of sensors increases and secondly the transmission i.e. the sensors may not be able to transmit at every time step, thus one may want to transmit after some local processing. A possible way to handle the first problem (reduce computational burden) is by using distributed Kalman filter [154].

Within this thesis when referring to a centralised approach direct fusion of all the ranges from all transmitter-receiver pairs is considered. In this case non-linear observation models should be used since the observations and target states are defined in different subspaces for e.g. polar and Cartesian coordinates. When referring to distributed fusion, local location estimates provided by each sensor of the network are fused. In this case linear models can be used since both observations and target states are defined in the same space i.e. both use for Cartesian coordinates.

APPENDIX V

OSPA METRIC

For performance evaluation of multiple target tracking problems a suitable distance measure is required which takes into consideration the possibility of target track birth/death and misdetection. Within this thesis target states are represented as a finite state set. Thus a suitable error metric would be able to compare two sets. Originally, the Hausdorff metric was considered for performance evaluation of multiple target tracking problems, however it was found to be insensitive to the difference between the number of elements of two sets [155]. The suggested W_p distance in [155] only partially solved this problem since it was not defined if one of the sets was empty and was inconsistent when the sets have different number of elements. Schuhmacher et al proposed a new metric called optimal subpattern assignment (OSPA) in [138] based on a metric between distributions of point processes. The proposed metric is based on the Wasserstein distance. The OSPA metric represents the sum of localisation (spatial distance) and cardinality distance via a cut-off parameter c. The localisation error is represented as the minimum sum of distances. The cut-off parameter c deals with any mismatch between the number of elements of the two finite sets being compared (the ground truth and estimated target states). The unassigned elements, extra target states representing either a missing existing target or an extra false target in the set of estimated targets, are penalised with maximum error c, which is considered to be the cardinality error. The maximum distance

between two elements of the sets being compared should be less than the cut-off value c. Let $X = \{x_1, x_2, ..., x_m\}$ and $Y = \{y_1, y_2, ..., y_n\}$ represent two finite sets. The OSPA metric of order p with cut-off c can be defined as

$$\bar{d}_p^{(c)}(X,Y) = \left(\frac{1}{n} \left(\min_{\pi \in \Pi_n} \sum_{i=1}^{m} d^{(c)}\big(x_i, y_{\pi(i)}\big)^p + c^p(n-m) \right) \right)^{1/p} \tag{V.1}$$

for $m \leq n$ where $d^{(c)}(x,y) = \min(c, ||x-y||)$ defines the distance between state x and state y for a cut-off $c > 0$. Π_k defines the set of permutations on $\{1, 2, ..., k\}$ for any $k \in \mathbb{N} = \{1, 2, ...\}$.

For $m > n$, $\bar{d}_p^{(c)}(X,Y) = \bar{d}_p^{(c)}(Y,X)$ and for $m = n = 0$ the distance is set to zero, i.e. $\bar{d}_p^{(c)}(X,Y) = \bar{d}_p^{(c)}(Y,X) = 0$.

The practical use of the OSPA metric is as follows (assume $m \leq n$):

- Find the m-point subpattern of Y (subset of Y consisting of m elements) closest to X in terms of the pth order Wasserstein distance metric.
- For each element y_j of Y, define α_j where $\alpha_j = c$ if there is no element from X assigned to it, or $\alpha_j = \min(c, ||y_j - x_i||)$ where x_i is the assigned element in X.
- Compute the pth order average $((1/n)\sum_{j=1}^{n} \alpha_j^p)^{1/p}$ for the defined α_j values.

The order $p \geq 1$ is used to define the order for the per-target location error and the order for the per-target cardinality error. These error component for $p < \infty$ and $m \leq n$ are given by

$$\begin{aligned} \bar{e}_{p,loc}^{(c)}(X,Y) &= \left(\frac{1}{n} \min_{\pi \in \Pi_n} \sum_{i=1}^{m} d^{(c)}\left(x_i, y_{\pi(i)}\right) \right)^{1/p} \\ \bar{e}_{p,card}^{(c)}(X,Y) &= \left(\frac{c^p(n-m)}{n} \right)^{1/p} \end{aligned} \tag{V.2}$$

The value p determines the sensitivity of the OSPA metric to outlier estimates. For $p = 1$ the OSPA metric measures the first order per-object error. The sum of the localisation and cardinality components is the total metric and the components of the metric are easier to interpret. This is the reason it is also used for the performance analysis within this thesis. An order of $p = 2$ would result in a more smooth distance curve and might be more appropriate to use. However the metric components of localisation and cardinality error can not be as easily interpreted.

The cut-off c determines the weighting of the cardinality component against the localisation error. Typically a cut-off value should be chosen which corresponds to the magnitude of a typical localisation error. If c corresponds to the magnitude of a typical localisation error, it is considered small and emphasises localisation errors. If c corresponds to the maximal distance between objects, it can be considered large and emphasises cardinality errors. The cut-off c can be interpreted as the penalty given if there is a false or missing estimate. Within this thesis cut-off value $c = 10$ is used. This means that if no targets are estimated and there are two targets present in the ground truth finite set, the OSPA metric has the value 10. If one of those targets is correctly estimated with small localisation error ε and the other is not estimated at all, the OSPA metric would have the value $5 + \varepsilon$.

Since the OSPA metric compares two finite sets composed of a finite collection of states, it was considered insufficient for performance analysis of multiple target tracking problems. For sufficiently analysing tracking problems the finite sets should be collections of sequences of target states. In this case target trajectories instead of target states would be compared. Different extensions have been proposed, such as the one in [156], however for the performance analysis in this thesis the OSPA metric was deemed sufficient and used.

BIBLIOGRAPHY

Own References

[1] S. Jovanoska, R. Zetik, R. Thomä, F. Govaers, K. Wild, and W. Koch, "Device-free indoor localization using a distributed network of autonomous UWB sensor nodes," in *Workshop on Sensor Data Fusion: Trends, Solutions, Applications (SDF)*, Oct 2013, pp. 1–6.

[2] S. Jovanoska and R. Thomä, "Multiple target tracking by a distributed UWB sensor network based on the PHD filter," in *15th International Conference on Information Fusion (FUSION)*, 2012, pp. 1095–1102.

[3] R. Zetik, H. Yan, E. Malz, S. Jovanoska, G. Shen, R. S. Thomä, R. Salman, T. Schultze, R. Tobera, H.-I. Willms, L. Reichardt, M. Janson, T. Zwick, W. Wiesbeck, T. Deißler, and J. Thielecke, *Ultra-Wideband Radio Technologies for Communications, Localization and Sensor Applications*. InTech, 2013, ch. Cooperative Localization and Object Recognition in Autonomous UWB Sensor Networks, pp. 179–240.

[4] G. Shen, R. Zetik, H. Yan, S. Jovanoska, E. Malz, and R. S. Thomä, "UWB localization of active nodes in realistic indoor environments," in *Loughborough Antennas & Propagation Conference (LAPC)*, Loughborough, UK, November 2011.

[5] G. Shen, R. Zetik, H. Yan, S. Jovanoska, and R. S. Thomä, "Localization of active UWB sensor nodes in multipath and NLOS environments," in *5th European Conference on Antennas and Propagation (EUCAP)*, Rome, Italy, April 2011.

[6] S. Jovanoska and R. Thomä, "Localization and tracking of multiple passive targets using distributed UWB sensor network," Technische Universität Ilmenau, Abschlussbericht zum Forschungskooperationvertrag zwischen Fraunhofer-FKIE und TU Ilmenau (FG EMT), 2012.

[7] S. Jovanoska, R. Zetik, F. Govaers, W. Koch, and R. Thomä, "Localization of multiple persons using UWB in indoor scenarios," 11th COST IC1004 Management Committee Meeting, Technische Universität Ilmenau in cooperation with Fraunhofer FKIE, Krakow, Poland, TD (14)11014, Sep 2014.

[8] S. Jovanoska, F. Govaers, and R. Thomä, *Novel Radar Techniques and Applications Volume 2: Waveform Diversity and Cognitive Radar, and Target Tracking and Data Fusion.* Institution of Engineering and Technology, 2017, vol. 2, ch. Person Tracking and Data Fusion for UWB Radar Applications, pp. 429–455.

[9] S. Jovanoska, F. Govaers, R. Thomä, and W. Koch, "Dynamic occlusion handling in the PHD filter for range-only tracking: Proof of concept," in *Workshop on Sensor Data Fusion: Trends, Solutions, Applications (SDF)*, Oct 2014, pp. 1–6.

[10] ——, "Dynamic-occlusion likelihood incorporation in a PHD filter based range-only tracking system," in *18th International Conference on Information Fusion (FUSION)*, Jul 2015, pp. 1078–1084.

[11] R. Zetik, S. Jovanoska, and R. Thomä, "Simple method for localisation of multiple tag-free targets using UWB sensor network," in *IEEE International Conference on Ultra-Wideband (ICUWB)*, Sep. 2011, pp. 268 –272.

[12] H. Yan, G. Shen, R. Zetik, S. Jovanoska, E. Malz, and R. S. Thomä, "Stationary symmetric object detection in unknown indoor environments," in *Loughborough Antennas & Propagation Conference (LAPC)*, Loughborough, UK, November 2011.

[13] R. Zetik, M. Eschrich, S. Jovanoska, and R. Thoma, "Looking behind a corner using multipath-exploiting UWB radar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, no. 3, pp. 1916–1926, July 2015.

[14] S. Jovanoska, M. Brötje, and W. Koch, "Multisensor data fusion for UAV detection and tracking," in *19th International Radar Symposium (IRS)*, June 2018, pp. 1–10.

References by Other Authors

[15] Y. Kilic, H. Wymeersch, A. Meijerink, M. Bentum, and W. Scanlon, "Device-free person detection and ranging in UWB networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 1, pp. 43–54, Feb 2014.

[16] T. Teixeira, G. Dublon, and A. Savvides, "A survey of human-sensing: Methods for detecting presence, count, location, track, and identity," *ENALAB technical report*, 2010.

[17] G. Deak, K. Curran, and J. Condell, "A survey of active and passive indoor localisation systems," *Computer Communications*, vol. 35, no. 16, pp. 1939 – 1954, 2012.

[18] C. Debono and E. Sammut, "Location estimation of an intruder in wireless ad hoc networks," in *Proc. 2008 IEEE Mediterranean Electrotechnical Conference*, 2008, pp. 158–162.

[19] A. Ko, H. Y. K. Lau, and N. M. Y. Lee, "Ais based distributed wireless sensor network for mobile search and rescue robot tracking," *J. Math. Model. Algorithms*, vol. 8, no. 2, pp. 227–242, 2009.

[20] A. Ko and H. Y. K. Lau, "Robot assisted emergency search and rescue system with a wireless sensor network," *International Journal of Advanced Science and Technology*, vol. 3, pp. 69–78, 2009.

[21] O. Sisma, A. Gaugue, C. Liebe, and J.-M. Ogier, "UWB radar: vision through a wall," *Telecommunication Systems*, vol. 38, no. 1-2, pp. 53–59, 2008. [Online]. Available: http://dx.doi.org/10.1007/s11235-008-9087-z

[22] M. Youssef, M. Mah, and A. Agrawala, "Challenges: Device-free passive localization for wireless environments," in *Proceedings of the 13th Annual ACM International Conference on Mobile Computing and Networking*. ACM, 2007, pp. 222–229.

[23] D. Zhang, J. Ma, Q. Chen, and L. Ni, "An RF-based system for tracking transceiver-free objects," in *5th Annual IEEE International Conference on Pervasive Computing and Communications*, March 2007, pp. 135–144.

[24] C. Xu, B. Firner, Y. Zhang, R. Howard, J. Li, and X. Lin, "Improving RF-based device-free passive localization in cluttered indoor environ-

ments through probabilistic classification methods," in *Proceedings of the 11th International Conference on Information Processing in Sensor Networks*. ACM, 2012, pp. 209–220.

[25] S. Nannuru, Y. Li, Y. Zeng, M. Coates, and B. Yang, "Radio frequency tomography for passive indoor multi-target tracking," *IEEE Transactions on Mobile Computing*, vol. 99, 2012.

[26] R. S. Thomä, O. Hirsch, J. Sachs, and R. Zetik, "UWB sensor networks for position location and imaging of objects and environments," in *The 2nd European Conference on Antennas and Propagation (EUCAP)*, Nov 2007, pp. 1–9.

[27] R. Zetik, J. Sachs, and P. Peyerl, "Through-wall imaging by means of UWB radar," in *Ultra-Wideband, Short-Pulse Electromagnetics 7*, F. Sabath, E. L. Mokole, U. Schenk, and D. Nitsch, Eds. Springer New York, 2007, pp. 613–622.

[28] M. Kmec, J. Sachs, P. Peyerl, P. Rauschenbach, R. Thomä, and R. Zetik, "A novel ultra-wideband real-time MIMO channel sounder architecture," in *XXVIIIth URSI General Assembly*, New Delhi, India, 2005.

[29] C. Falsi, D. Dardari, L. Mucchi, and M. Z. Win, "Time of arrival estimation for UWB localizers in realistic environments," *EURASIP Journal on Applied Signal Processing*, pp. 1–13, 2006.

[30] D. B. Jourdan, D. Dardari, and M. Z. Win, "Position error bound for UWB localization in dense cluttered environments," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 44, no. 2, pp. 613–628, April 2008.

[31] D. Dardari, A. Conti, U. Ferner, A. Giorgetti, and M. Z. Win, "Ranging with ultrawide bandwidth signals in multipath environments," *Proceedings of the IEEE*, vol. 97, no. 2, pp. 404–426, Feb 2009.

[32] S. Gezici, Z. Tian, G. Giannakis, H. Kobayashi, A. Molisch, H. Poor, and Z. Sahinoglu, "Localization via ultra-wideband radios: a look at positioning aspects for future sensor networks," *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 70–84, Jul. 2005.

[33] H. Saghir, M. Heddebaut, F. Elbahhar, J.-M. Rouvaen, A. Menhaj-Rivenq, and J.-P. Ghys, "Train-to-wayside wireless communication in tunnel using ultra-wide-band and time reversal," *Transportation*

Research Part C: Emerging Technologies, vol. 17, no. 1, pp. 81 – 97, 2009.

[34] H. Nikookar and R. Prasad, *Introduction to Ultra Wideband for Wireless Communications*, 1st ed., ser. Signals and Communication Technology. Springer Science + Business Media B.V., 2009.

[35] C. N. Paulson, J. T. Chang, C. E. Romero, J. Watson, F. J. Pearce, and N. Levin, "Ultra-wideband radar methods and techniques of medical sensing and imaging," in *Smart Medical and Biomedical Sensor Technology III*, ser. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, B. M. Cullum & J. C. Carter, Ed., vol. 6007, Nov. 2005, pp. 96–107.

[36] R. Zetik, J. Sachs, and R. S. Thoma, "UWB short-range radar sensing - the architecture of a baseband, pseudo-noise UWB radar sensor," *IEEE Instrumentation Measurement Magazine*, vol. 10, no. 2, pp. 39–45, April 2007.

[37] R. Herrmann, "M-sequence based ultra-wideband radar and its application to crack detection in salt mines," Ph.D. dissertation, Technische Universität Ilmenau, 2011.

[38] J. Teizer, M. Venugopal, and A. Walia, "Ultrawideband for automated real-time three-dimensional location sensing for workforce, equipment, and material positioning and tracking," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2081, pp. 56–64, 2008.

[39] J. Sachs, M. Helbig, R. Herrmann, M. Kmec, K. Schilling, E. Zaikov, and P. Rauschenbach, "Trapped victim detection by pseudo-noise radar," in *Proceedings of the 1st International Conference on Wireless Technologies for Humanitarian Relief*. ACM, 2011, pp. 265–272.

[40] J. Sachs, P. Peyerl, R. Zetik, and S. Crabbe, "M-sequence ultra-wideband-radar: state of development and applications," in *Proceedings of the International Radar Conference*, September 2003, pp. 224–229.

[41] A. Giretti, A. Carbonari, B. Naticchia, and M. DeGrassi, "Design and first development of an automated real-time safety management system for construction sites," *Journal of Civil Engineering and Management*, vol. 15, no. 4, pp. 325–336, 2009.

[42] B. Anderson and J. Moore, *Optimal Filtering.* Englewood Cliffs, NJ: Prentice-Hall, 1979.

[43] S. Särkkä, *Bayesian Filtering and Smoothing*, ser. Institute of Mathematical Statistics Textbooks. Cambridge University Press, 2013.

[44] T. Lefebvre, H. Bruyninckx, and J. D. Schutter, "Kalman filters for non-linear systems: a comparison of performance," *International Journal of Control*, vol. 77, no. 7, pp. 639–653, 2004.

[45] D. Simon, *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches.* New York, NY, USA: Wiley-Interscience, 2006.

[46] H. W. Sorenson and A. R. Stubberud, "Recursive filtering for systems with small but non-negligible non-linearities," *International Journal of Control*, vol. 7, no. 3, pp. 271–280, 1968.

[47] S. J. Julier, J. K. Uhlmann, and H. F. Durrant-Whyte, "A new approach for filtering nonlinear systems," in *Proceedings of 1995 American Control Conference*, vol. 3, June 1995, pp. 1628–1632 vol.3.

[48] D. Alspach and H. Sorenson, "Nonlinear bayesian estimation using gaussian sum approximations," *IEEE Transactions on Automatic Control*, vol. 17, no. 4, pp. 439–448, August 1972.

[49] Y. Bar-Shalom and T. E. Fortmann, *Tracking and Data Association.* New York: Academic Press, 1988.

[50] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems.* Norwood, MA: Artech House, 1999.

[51] G. W. Pulford, "Taxonomy of multiple target tracking methods," *IEE Proceedings - Radar, Sonar and Navigation*, vol. 152, no. 5, pp. 291–304, Oct 2005.

[52] G. van Keuk, "Multihypothesis tracking with electronically scanned radar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 31, no. 3, pp. 916–927, Jul. 1995.

[53] R. Mahler, "Multitarget Bayes filtering via first-order multitarget moments," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, no. 4, pp. 1152–1178, Oct. 2003.

[54] B.-N. V. Vo and W.-K. Ma, "The Gaussian mixture probability hypothesis density filter," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4091–4104, Nov. 2006.

[55] B.-N. Vo, S. Singh, and A. Doucet, "Sequential monte carlo methods for multitarget filtering with random finite sets," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 41, no. 4, pp. 1224–1245, Oct. 2005.

[56] R. Streit, "A technique for deriving multitarget intensity filters using ordinary derivatives," *Journal of Advances in Information Fusion*, vol. 9, pp. 3–12, 2014.

[57] C. Degen, "Finite point processes and their application to target tracking," Ph.D. dissertation, Rheinische Friedrich-Wilhelms-Universität Bonn, 2017.

[58] S. Gezici, "A survey on wireless position estimation," *Wireless Personal Communications*, vol. 44, pp. 263–282, Feb 2008.

[59] Y. Qi, H. Kobayashi, and H. Suda, "Analysis of wireless geolocation in a non-line-of-sight environment," *IEEE Transactions on Wireless Communications*, vol. 5, no. 3, pp. 672–681, March 2006.

[60] Ilmsens GmbH. Last accessed: 31.03.2018. [Online]. Available: www.uwb-shop.com

[61] J. Sachs, *Handbook of Ultra-Wideband Short-Range Sensing:Theory, Sensors, Applications*, 1st ed. Wiley-VCH Verlag GmbH & Co. KGaA, 2012.

[62] T. Kaiser, C. Senger, A. Eltaher, and B. T. Sieskul, *Ultra-Wideband: Antennas and Propagation for Communications, Radar and Imaging.* Chichester, UK.: John Wiley & Sons, Ltd, NOV 2006, ch. Localisation in NLOS Scenarios with UWB Antenna Arrays, pp. 389–411.

[63] F. Thiel, M. Helbig, U. Schwarz, C. Geyer, G. Rimkus, W. A. Kaiser, I. Hilger, M. Hein, J. Sachs, and F. Seifert, "Implementation of ultra-wideband sensors for biomedical applications," *Frequenz, Journal of RF Engineering and Telecommunucations*, vol. 63, no. 9-10, pp. 221–224, Oct 2009.

[64] R. Zetik, M. Kmec, J. Sachs, and R. S. Thomä, "Real-time MIMO channel sounder for emulation of distributed ultrawideband systems," *International Journal of Antennas and Propagation*, vol. 2014, 2014, 16 pages.

[65] Z. Sahinoglu, S. Gezici, and I. Güvenc, *Ultra-wideband Position-*

ing Systems: Theoretical Limits, Ranging Algorithms, and Protocols. Cambridge University Press, Oct 2008.

[66] P. Swerling, "Probability of detection for fluctuating targets," *IRE Transactions on Information Theory*, vol. 6, no. 2, pp. 269–308, April 1960.

[67] ——, "Radar probability of detection for some additional fluctuating target cases," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 33, no. 2, pp. 698–709, April 1997.

[68] A. Farina, A. Russo, and F. Studer, "Coherent radar detection in log-normal clutter," *IEE Proceedings F - Communications, Radar and Signal Processing*, vol. 133, no. 1, pp. 39–53, Feb 1986.

[69] A. Farina, A. Russo, F. Scannapieco, and S. Barbarossa, "Theory of radar detection in coherent weibull clutter," *IEE Proceedings F - Communications, Radar and Signal Processing*, vol. 134, no. 2, pp. 174–190, Apr 1987.

[70] M. I. Skolnik, *Introduction to Radar Systems*, 2nd ed. New York, NY: McGraw-Hill, 1980.

[71] M. A. Richards, *Fundamentals of Radar Signal Processing*, ser. Electronic Engineering. New York, NY: McGraw-Hill, 2005.

[72] A. Safaai-Jazi, S. M. Riad, A. Muqaibel, and A. Bayram, "Ultra-wideband propagation measurements and channel modeling," in *Report on Through-the-Wall Propagation and Material Characterization*, Nov. 2002, DARPA NETEX Program Report.

[73] E. Paolini, A. Giorgetti, M. Chiani, R. Minutolo, and M. Montanari, "Localization capability of cooperative anti-intruder radar systems," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, no. 1, p. 726854, 2008.

[74] S. Bartoletti, A. Conti, and A. Giorgetti, "Analysis of UWB radar sensor networks," in *IEEE International Conference on Communications (ICC)*, May 2010, pp. 1–6.

[75] M. I. Skolnik, "An analysis of bistatic radar," *IRE Transactions on Aerospace and Navigational Electronics*, vol. ANE-8, no. 1, pp. 19–27, March 1961.

[76] S. Venkatesh, C. Anderson, N. Rivera, and R. Buehrer, "Implementation and analysis of respiration-rate estimation using impulse-based

UWB," in *IEEE Military Communications Conference*, Oct 2005, pp. 3314–3320 Vol. 5.

[77] N. Rivera, S. Venkatesh, C. Anderson, and R. Buehrer, "Multi-target estimation of heart and respiration rates using ultra wide-band sensors," in *Proceedings of the 14th European Signal Processing Conference (EUSIPCO)*, 2006.

[78] A. G. Yarovoy, L. P. Ligthart, J. Matuzas, and B. Levitas, "UWB radar for human being detection," *IEEE Aerospace and Electronic Systems Magazine*, vol. 21, no. 11, pp. 22–26, Nov 2006.

[79] G. Ossberger, T. Buchegger, E. Schimback, A. Stelzer, and R. Weigel, "Non-invasive respiratory movement detection and monitoring of hidden humans using ultra wideband pulse radar," in *International Workshop on Ultra Wideband Systems Joint with Conference on Ultra Wideband Systems and Technologies. Joint UWBST IWUWBS 2004 (IEEE Cat. No.04EX812)*, May 2004, pp. 395–399.

[80] Y. Xu, S. Wu, J. Shao, J. Chen, and G. Fang, "Life detection and location by MIMO ultra wideband radar," in *14th International Conference on Ground Penetrating Radar (GPR)*, June 2012, pp. 80–84.

[81] S. Gezici, "Theoretical limits for estimation of periodic movements in pulse-based UWB systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 3, pp. 405–417, Oct 2007.

[82] J. Li, Z. Zeng, J. Sun, and F. Liu, "Through-wall detection of human being's movement by UWB radar," *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 6, pp. 1079–1083, Nov 2012.

[83] V. Casadei, N. Nanna, and D. Dardari, "Experimental study in breath detection and human target ranging in the presence of obstacles using ultra-wideband signals," *International Journal of Ultra Wideband Communications and Systems*, vol. 2, no. 2, pp. 116–123, 2011.

[84] A. Nezirovic, S. Tesfay, A. Valavan, and A. Yarovoy, "Experimental study on human breathing cross section using UWB impulse radar," in *European Radar Conference (EuRAD)*, Oct 2008, pp. 1–4.

[85] S.-c. S. Cheung and C. Kamath, "Robust techniques for background subtraction in urban traffic video," in *SPIE Visual Communications and Image Processing*, vol. 5308, Jan. 2004, pp. 881–892.

[86] D. J. Daniels, Ed., *Ground penetrating radar*, 2nd ed., ser. IET

Radar, Sonar, Navigation and Avionics. London, UK: Institution of Engineering and Technology, 2004, vol. 15.

[87] M. Piccardi, "Background subtraction techniques: a review," in *IEEE International Conference on Systems, Man and Cybernetics*, vol. 4, Oct 2004, pp. 3099–3104.

[88] R. Zetik, S. Crabbe, J. Krajnak, P. Peyerl, J. Sachs, and R. Thomä, "Detection and localization of persons behind obstacles using m-sequence through-the-wall radar," in *Proceedings of the SPIE*, vol. 6201, Apr 2006, pp. 62 010I–62 010I–12.

[89] D. Dardari, C.-C. Chong, and M. Z. Win, "Analysis of threshold-based TOA estimator in UWB channels," in *Proceedings of European Signal Processing Conference (EUSIPCO)*, Sep 2006.

[90] I. Guvenc and Z. Sahinoglu, "Threshold-based TOA estimation for impulse radio UWB systems," in *Proceedings of IEEE International Conference on Ultra-Wideband (ICUWB)*, Sep. 2005, pp. 420–425.

[91] D. Dardari, C.-C. Chong, and M. Z. Win, "Threshold-based time-of-arrival estimators in UWB dense multipath channels," *IEEE Transactions on Communications*, vol. 56, no. 8, pp. 1366–1378, Aug. 2008.

[92] G. Van Der Spek, "Detection of a distributed target," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-7, no. 5, pp. 922–931, Sep 1971.

[93] J. D. Taylor, *Ultra-wideband Radar Technology*. CRC Press, 2002.

[94] G. Minkler and J. Minkler, "CFAR: The principles of automatic radar detection in clutter," *NASA STI/Recon Technical Report A*, vol. 90, p. 23371, 1990.

[95] P. Dutta, A. Arora, and S. Bibyk, "Towards radar-enabled sensor networks," in *Proceedings of the 5th International Conference on Information Processing in Sensor Networks*, Apr. 2006, pp. 467 –474.

[96] S. C. Johnson, "Hierarchical clustering schemes," *Psychometrika*, vol. 2, pp. 241–254, 1967.

[97] J. Rovnakova and D. Kocur, "Compensation of wall effect for through wall tracking of moving targets," *Radioengineering*, vol. 18, no. 2, pp. 189–195, June 2009.

[98] M. Daun and C. R. Berger, "Track initialization in a multistatic

DAB/DVB-T network," in *11th International Conference on Information Fusion (FUSION)*, June 2008, pp. 1–8.

[99] Y. Zhou, C. L. Law, Y. L. Guan, and F. Chin, "Localization of passive target based on UWB backscattering range measurement," in *IEEE International Conference on Ultra-Wideband (ICUWB)*, Sep. 2009, pp. 145 –149.

[100] M. Svecova and, D. Kocur, R. Zetik, and J. Rovnnakova, "Target localization by a multistatic UWB radar," in *20th International Conference Radioelektronika*, Apr. 2010, pp. 1 –4.

[101] X. Zhao, A. Gaugue, C. Lièbe, J. Khamlichi, and M. Ménard, "Through the wall detection and localization of a moving target with a bistatic UWB radar system," in *European Radar Conference*, Oct. 2010, pp. 204 –207.

[102] J. Rovnakova and D. Kocur, "Data fusion from UWB radar network: Preliminary experimental results," in *21st International Conference Radioelektronika*, Apr. 2011, pp. 1 –4.

[103] S. Chang, M. T. Wolf, and J. W. Burdick, "Human detection and tracking via ultra-wideband (UWB) radar," in *IEEE International Conference on Robotics and Automation*, May 2010, pp. 452–457.

[104] Y. He, T. Savelyev, and A. Yarovoy, "Two-stage algorithm for extended target tracking by multistatic UWB radar," in *IEEE CIE International Conference on Radar*, vol. 1, 2011, pp. 795–799.

[105] J. Shen, A. Molisch, and J. Salmi, "Accurate passive location estimation using TOA measurements," *IEEE Transactions on Wireless Communications,*, vol. 11, no. 6, pp. 2182–2192, 2012.

[106] S. Chang, R. Sharan, M. Wolf, N. Mitsumoto, and J. Burdick, "UWB radar-based human target tracking," in *IEEE Radar Conference*, May 2009, pp. 1–6.

[107] S. Chang, N. Mitsumoto, and J. Burdick, "An algorithm for UWB radar-based human detection," in *IEEE Radar Conference*, May 2009, pp. 1–6.

[108] K. Yu, J. P. Montillet, A. Rabbachin, P. Cheong, and I. Oppermann, "UWB location and tracking for wireless embedded networks," *Signal Processing*, vol. 86, no. 9, pp. 2153 – 2171, 2006, Special Section: Signal Processing in UWB Communications.

[109] R. Thomä, O. Hirsch, J. Sachs, and R. Zetik, "UWB sensor networks for reconnaissance of environments," Technische Universität Ilmenau, FKIE technical report, 2009.

[110] S. Bartoletti, A. Conti, A. Giorgetti, and M. Win, "Sensor radar networks for indoor tracking," *IEEE Wireless Communications Letters*, vol. 3, no. 2, pp. 157–160, Apr 2014.

[111] C. Chang and A. Sahai, "Object tracking in a 2D UWB sensor network," in *Conference Record of the 38th Asilomar Conference on Signals, Systems and Computers*, vol. 1, Nov 2004, pp. 1252–1256.

[112] M. Chiani, A. Giorgetti, M. Mazzotti, R. Minutolo, and E. Paolini, "Target detection metrics and tracking for UWB radar sensor networks," in *IEEE International Conference on Ultra-Wideband (ICUWB)*, Sept 2009, pp. 469–474.

[113] M. Arik and O. Akan, "Collaborative mobile target imaging in UWB wireless radar sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 6, pp. 950–961, Aug. 2010.

[114] G. Shen, "Localization of active nodes within distributed ultra-wideband sensor networks in multipath environments," Ph.D. dissertation, Technische Universität Ilmenau, 2012.

[115] S. Cheshire and D. Steinberg, *Zero Configuration Networking: The Definitive Guide*, ser. Definitive Guide Series. O'Reilly Media, 2006.

[116] ZeroC Inc. The internet communication engine. Last accessed: 04.10.2018. [Online]. Available: https://zeroc.com/products/ice

[117] D. Kocur, M. Svecova, and J. Rovnakova, "Through-the-wall localization of a moving target by two independend ultra wideband (UWB) radar systems," *Sensors*, vol. 13, pp. 11 969–11 997, 2013.

[118] V. Algeier, R. S. Thomä, and W. Koch, "Robust likelihood function for blind mobile terminal localization," in *GI Jahrestagung*, 2010, pp. 850–855.

[119] V. Algeier, "Blind localization of mobile terminals in urban scenarios," Ph.D. dissertation, Technische Universität Ilmenau, 2010.

[120] A. N. Steinberg, C. L. Bowman, and F. E. White, "Revisions to the JDL data fusion model," in *Proceedings of the SPIE Sensor Fusion: Architectures, Algorithms, and Applications*, vol. 3719, 1999.

[121] A. Jazwinski, *Stochastic processes and filtering theory*, 1st ed., ser. Mathematics in science and engineering. Academic Press, 1970, vol. 64.

[122] S. J. Julier and J. K. Uhlmann, "New extension of the kalman filter to nonlinear systems," in *Proceedings of the SPIE Signal Processing, Sensor Fusion, and Target Recognition VI*, vol. 3068, 1997.

[123] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and Computing*, vol. 10, no. 3, pp. 197–208, 2000.

[124] N. Gordon, D. Salmond, and A. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," *IEE Proceedings F Radar and Signal Processing*, vol. 140, no. 2, pp. 107–113, 1993.

[125] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Norwood, MA: Artech House, 2004.

[126] N. Gordon, D. Salmond, and D. Fisher, "Bayesian target tracking after group pattern distortion," in *Proceedings of the SPIE*, 1997.

[127] S. Blackman, *Multiple-Target Tracking with Radar Applications*. Norwood, MA: Artech House, 1986.

[128] Y. Bar-Shalom and X. Rong Li, *Multitarget-Multisensor Tracking: principles and techniques*. Storrs, CT: YBS Publishing, 1995.

[129] D. B. Reid, "An algorithm for tracking multiple targets," *IEEE Transactions on Automatic Control*, vol. 24, no. 6, pp. 843–854, 1979.

[130] K. Panta, D. E. Clark, and B. N. Vo, "Data association and track management for the gaussian mixture probability hypothesis density filter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 3, pp. 1003–1016, July 2009.

[131] L. Lin, Y. Bar-Shalom, and T. Kirubarajan, "Track labeling and PHD filter for multitarget tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, no. 3, pp. 778 –795, Jul. 2006.

[132] B. Ristic, D. Clark, B.-N. Vo, and B.-T. Vo, "Adaptive target birth intensity for PHD and CPHD filters," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 2, pp. 1656–1668, Apr 2012.

[133] R. Mahler, *Advances in Statistical Multisource-multitarget Information Fusion*. Artech House, 2014, ch. Multisensor PHD and CPHD Filters.

[134] N. T. Pham, W. Huang, and S. Ong, "Multiple sensor multiple object tracking with GMPHD filter," in *10th International Conference on Information Fusion*, Jul 2007, pp. 1–7.

[135] Y. Bar-Shalom, X. Rong Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*. New York: John Wiley & Sons, 2001.

[136] X. R. Li and V. P. Jilkov, "Survey of maneuvering target tracking. part i. dynamic models," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, no. 4, pp. 1333–1364, Oct 2003.

[137] E. Mazor, A. Averbuch, Y. Bar-Shalom, and J. Dayan, "Interacting multiple model methods in target tracking: a survey," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 34, no. 1, pp. 103–123, Jan. 1998.

[138] D. Schuhmacher, B.-T. Vo, and B.-N. Vo, "A consistent metric for performance evaluation of multi-object filters," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3447–3457, Aug 2008.

[139] S. Gezici, Z. Sahinoglu, H. Kobayashi, and H. V. Poor, *Ultra Wideband Wireless Communication*. Hoboken, NJ, USA.: John Wiley & Sons, Inc., 2005, ch. Ultra Wideband Geolocation, pp. 43–75.

[140] D. Urdzik, R. Zetik, D. Kocur, and J. Rovnakova, "Shadowing effect investigation for the purposes of person detection and tracking by UWB radars," in *The 7th German Microwave Conference*, 2012, pp. 1–4.

[141] D. Kocur, J. Fortes, and M. Švecová, "Multiple moving person tracking by UWB sensors: the effect of mutual shielding persons and methods reducing its impacts," *EURASIP Journal on Wireless Communications and Networking*, vol. 2017, no. 1, p. 68, Apr 2017.

[142] J. Hoffman, M. Spranger, D. Gohring, and M. Jungel, "Making use of what you don't see: negative information in Markov localization," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Aug 2005, pp. 2947–2952.

[143] S. Thrun, W. Burgard, and D. Fox, *Probabilistic robotics*. MIT press, 2005.

[144] K. Tischler and H. Vogt, "A sensor data fusion approach for the integration of negative information," in *10th International Conference on Information Fusion*, Jul 2007, pp. 1–7.

[145] W. Koch, "On exploiting 'negative' sensor evidence for target tracking and sensor data fusion," *Information Fusion*, vol. 8, no. 1, pp. 28–39, 2007.

[146] M. Powers, R. Ravichandran, F. Dellaert, and T. Balch, "Improving multirobot multitarget tracking by communicating negative information," in *Multi-Robot Systems. From Swarms to Intelligent Automata Volume III*, L. Parker, F. Schneider, and A. Schultz, Eds. Springer Netherlands, 2005, pp. 107–117.

[147] A. Andriyenko, S. Roth, and K. Schindler, "An analytical formulation of global occlusion reasoning for multi-target tracking," in *IEEE International Conference on Computer Vision Workshops*, Nov 2011, pp. 1839–1846.

[148] K. Wyffels and M. Campbell, "Negative observations for multiple hypothesis tracking of dynamic extended objects," in *American Control Conference (ACC)*, June 2014, pp. 642–647.

[149] ——, "Negative information for occlusion reasoning in dynamic extended multiobject tracking," *IEEE Transactions on Robotics*, vol. 31, no. 2, pp. 425–442, April 2015.

[150] ——, "Modeling and fusing negative information for dynamic extended multi-object tracking," in *IEEE International Conference on Robotics and Automation*, May 2013, pp. 3176–3182.

[151] E. Zaikov, J. Sachs, M. Aftanas, and J. Rovnakova, "Detection of trapped people by UWB radar," in *German Microwave Conference (GeMIC)*, Mar. 2008, pp. 1 –4.

[152] J. Sachs and R. Hermann, "M-sequence-based ultra-wideband sensor network for vitality monitoring of elders at home," *IET Radar, Sonar & Navigation*, vol. 9, pp. 125–137(12), Feb 2015.

[153] M. Ruggiano and P. v. Genderen, "Wideband ambiguity function and optimized coded radar signals," in *2007 European Radar Conference*, Oct 2007, pp. 142–145.

[154] R. Olfati-Saber, "Distributed kalman filtering for sensor networks,"

in *46th IEEE Conference on Decision and Control*, Dec 2007, pp. 5492–5498.

[155] J. R. Hoffman and R. P. S. Mahler, "Multitarget miss distance via optimal assignment," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 34, no. 3, pp. 327–336, 2004.

[156] B. Ristic, B. Vo, D. Clark, and B. Vo, "A metric for performance evaluation of multi-target tracking algorithms," *IEEE Transactions on Signal Processing*, vol. 59, no. 7, pp. 3452–3457, 2011.