Improving Accuracy in Ultra-Wideband Indoor Position Tracking through Noise Modeling and Augmentation

---

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Electrical Engineering

---

by
Salil P. Banerjee
December 2012

Dr. Adam W. Hoover, Committee Chair
Dr. Richard R. Brooks
Dr. John N. Gowdy
Dr. Eric R. Muth

Abstract

The goal of this research is to improve the precision in tracking of an ultra-wideband (UWB) based Local Positioning System (LPS). This work is motivated by the approach taken to improve the accuracies in the Global Positioning System (GPS), through noise modeling and augmentation. Since UWB indoor position tracking is accomplished using methods similar to that of the GPS, the same two general approaches can be used to improve accuracy.

Trilateration calculations are affected by errors in distance measurements from the set of fixed points to the object of interest. When these errors are systemic, each distinct set of fixed points can be said to exhibit a unique set noise. For UWB indoor position tracking, the set of fixed points is a set of sensors measuring the distance to a tracked tag. In this work we develop a noise model for this sensor set noise, along with a particle filter that uses our set noise model. To the author's knowledge, this noise has not been identified and modeled for an LPS. We test our methods on a commercially available UWB system in a real world setting. From the results we observe approximately 15% improvement in accuracy over raw UWB measurements.

The UWB system is an example of an aided sensor since it requires a person to carry a device which continuously broadcasts its identity to determine its location. Therefore the location of each user is uniquely known even when there are multiple users present. However, it suffers from limited precision as compared to some un-

aided sensors such as a camera which typically are placed line of sight (LOS). An unaided system does not require active participation from people. Therefore it has more difficulty in uniquely identifying the location of each person when there are a large number of people present in the tracking area. Therefore we develop a generalized fusion framework to combine measurements from aided and unaided systems to improve the tracking precision of the aided system and solve data association issues in the unaided system. The framework uses a Kalman filter to fuse measurements from multiple sensors. We test our approach on two unaided sensor systems: Light Detection And Ranging (LADAR) and a camera system. Our study investigates the impact of increasing the number of people in an indoor environment on the accuracies using a proposed fusion framework. From the results we observed that depending on the type of unaided sensor system used for augmentation, the improvement in precision ranged from $6 - 25\%$ for up to 3 people.

Dedication

This work is dedicated to my loving parents and fiancée who have always believed in me.

Acknowledgments

First, I would like to thank God for guiding the path to my degree by providing me with wonderful parents, a great advisor, loving fiancée, relatives and good friends. Next, I would like to sincerely thank my advisor Dr. Adam Hoover for his foresight that has prevented me from making hasty decisions and his infinite patience when I made mistakes. He has spent countless hours discussing ideas, reviewing and critiquing my work, providing me with guidance when I needed it and encouraging me when the future seemed bleak. His attention to detail, perseverance and teachings has made me a better researcher. I would also like to thank other members of my committee, Dr. Richard Brooks, Dr. John Gowdy, and Dr. Eric Muth for their time spent in reviewing my work and providing valuable insight and positive guidance.

I sincerely thank Dr. Damon Woodard from the Department of Computer Science for his financial support and giving me an opportunity to work with him. He introduced me to the exciting field of biometrics. His guidance and feedback has been very valuable and I have learned a lot from him.

I cannot forget the contributions of Bill Suski, who helped jump start my research when I had hit mid-academic life crisis. The hours we spent collecting data, exchanging ideas and brainstorming techniques to solve issues related to our research were an enjoyable experience. My colleagues Weixin Wu and Yujie Dong have also helped me in various ways to achieve my goals.

I am thankful for wonderful friends that I found in Vidya, Shubhada, Trupti, Shivram, Vikram, Shweta, Vibhor, Sneha and others. The late night movies, outings, tennis, long discussions on research and life have made my time in Clemson enjoyable. A special thanks to Vidya, Shubhada and Trupti who have been like elder sisters and taken care of me. I thank my roommates Raul, César, Dorismel, and Felipe who have put up with my idiosyncrasies, especially during the last few months. I would also like to thank Elizabeth, Lane, David and John in the ECE department for their help during my academic life in Clemson.

I would like to thank my parents, Partha and Jaiothi Banerjee who have always strived hard to provide me with the best. It is due to their hard work, sacrifice, and love that I have been able to achieve my goals.

Finally, I would like to thank my fiancée, Divina who has been a constant source of motivation and encouragement. It was her love and belief that helped me sail through dark times of my life. She has been a ray of sunshine with her positive attitude and constant smile without which life would be dull and uninteresting. If not for your rock-solid support and love I would not have been able to complete my degree.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

In the past few decades, the precision of radionavigation aids [71] used for outdoor navigation such as Global Navigation Satellite Systems (GNSS) have improved from hundreds of meters to a few meters. This increase in accuracy has led to an expansion of the original applications in military vehicle tracking and battlespace awareness to include civilian vehicle navigation and automated farming [85]. The accuracy of GNSSs have been improved in two general ways: noise modeling and augmentations. Systematic modeling of noise due to atmospheric effects, timing jitter, and satellite constellation is typically used in a Kalman filtering framework along with related signal processing approaches [71, 106, 117] to reduce the noise and improve the precision. Augmentation methods include using additional signals such as in differential land-based signals or position estimates from other hardware such as inertial systems to improve the precision of raw GNSS measurements [52].

Indoor position tracking is a new area in navigation which facilitates navigation within a building-sized area. For example, navigating a large building such as an airport can be confusing to the unexperienced traveler. A navigational aid similar to one used for outdoor navigation could help in such a situation. However, current aids used

Acronym	Meaning
ADC	Analog to Digital Converter
AOA	Angle of Arrival
CLEAR	Classification of Events, Activities and Relationships
EKF	Extended Kalman Filter
ETISEO	Evaluation du Traitment et de L'Interpretation de Séquences Video
GLONASS	Globalnaya navigatsionnaya sputnikovaya sistema
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
IMU	Inertial Measurement Unit
IR	Infrared
LOS	Line of Sight
LPS	Local Positioning System
MEMS	Micro Electro Mechanical Sensors
MOT	Multi-Object Tracking
NAVSAT	Navigation Satellite System
NLOS	Non-Line of Sight
PDF	Probability Distribution Function
PSD	Power Spectral Density
RF	Radio Frequency
RFID	Radio Frequency Identification
TDOA	Time Difference of Arrival
TOA	Time of Arrival
USA	United States of America
UWB	Ultra-wideband
WLAN	Wireless Local Area Networks

Table 1.1: Table of acronyms.

for outdoor navigation cannot work indoors due to low signal strength. Technologies such as wireless local area networks (WLAN), radio frequency identification (RFID), cameras, and infrared (IR) have been used for indoor positioning. The accuracy of current positioning systems in small buildings range from 30-100 cm [81, 112, 140], suitable for applications requiring rough room-level precision such as location based navigation, surveillance [78], asset tracking [93, 102] and indoor vehicle positioning. UWB is a relatively new technology used for indoor position tracking which uses the principle of trilateration, similar to GNSSs. Improving the precision of current indoor positioning systems could expand potential applications to telepresence, augmented reality, military training [35, 65], entertainment and medicine [51, 58]. In this work, we are inspired by methods used to improve the precision of a GNSS through noise modeling and augmentation methods.

The following sections provide some background on GNSS, LPS, UWB technology, aspects of data fusion, sensor ontology, and filtering techniques. These ideas form the basis for the work undertaken for this dissertation.

1.1 Global Navigation Satellite Systems

A GNSS is a global system of satellites that provides geo-spatial tracking information. The GPS is the first GNSS developed by the United States. The idea of GPS was conceived in the early 1960s in order to improve the tracking accuracy of the existing TRANSIT, also known as NAVSAT (Navy Navigation Satellite System). It was primarily developed for military applications such as tracking intercontinental ballistic missiles, ships and battlespace awareness [87]. The project was renamed Navstar-GPS after other branches of the US military joined the project, but the shortened name GPS is popularly used. The GPS consists of a constellation of 24

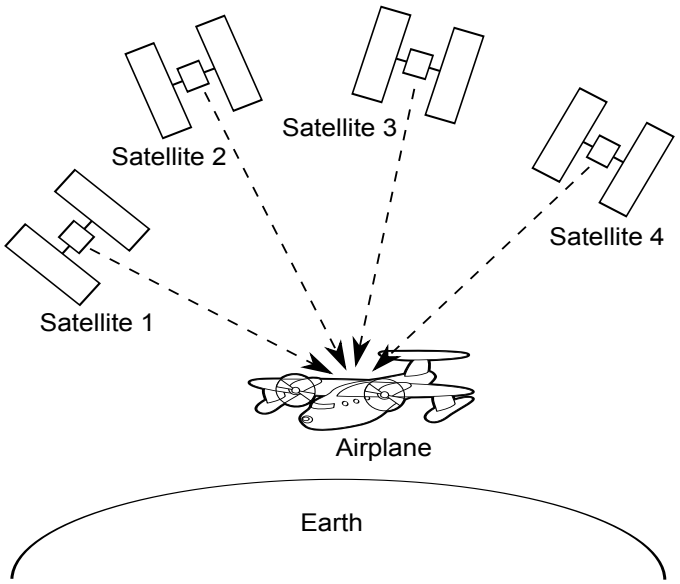


Figure 1.1: GNSS used for tracking an airplane.

evenly spaced satellites placed in 12 hour orbits and inclined at 55° orbit to the equatorial plane [62].

The Globalnaya Navigatsionnaya Sputnikovaya Sistema (GLONASS) is another operational GNSS developed by Russia. Two other GNSSs currently under development are the Galileo positioning system, a joint effort by the European Union and the European Space Agency, and the Compass (or Bediou-2) navigation system by China. Although GNSS signals are primarily used for outdoor tracking, a brief discussion of this technology is provided since some of the concepts used for UWB indoor tracking are similar [93].

The principle of a GNSS is to measure the time it takes for a satellite to broadcast the signal to a receiver [87]. This is known as measuring the TOA (time of arrival). Each satellite contains an atomic clock providing precise timing which is synchronized using ground control stations. At least four satellites are required to calculate the position of the receiver. Figure 1.1 shows a diagrammatic representation of four satellites being used to track an airplane in three dimensions. The satellites

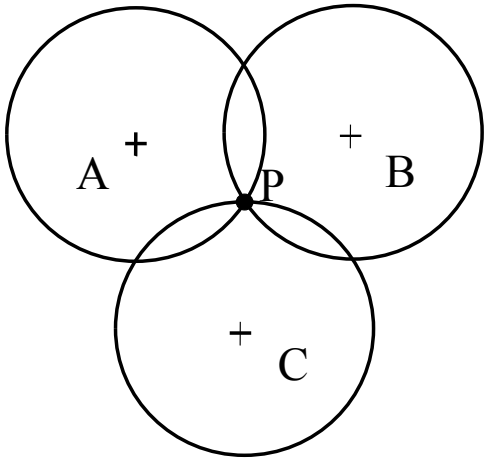


Figure 1.2: Trilateration in the presence of noiseless signals.

in a GNSS have been configured around the earth in such a manner that at any point of time there are at least 4 satellites overhead.

Figure 1.2 illustrates the technique used to calculate the position of the receiver from the satellite signals. This technique is known as *trilateration*. A single measurement places the receiver on a known circle (sphere in 3D) from the satellite. The '+' indicates the center where the satellite is placed while the solid circle indicates the possible locations of the receiver. Using the knowledge of the time taken to receive the signal and multiplying it by the speed of light gives the distance of the satellite from the receiver. Using at least three satellites, the 2D position of the receiver can be calculated by finding the intersection of the circles. Similarly, this idea can be expanded to 3D using the intersection of spheres and 4 satellites. There are several variations on this technique, including measuring the time difference of arrival (TDOA) and angle of arrival (AOA). Each of these methods varies somewhat in how each distance is measured, but the basic principle is the same.

In practice, the range measurement from each satellite to the receiver has an error in measuring the TOA of the signal which in turn affects the accuracy of the

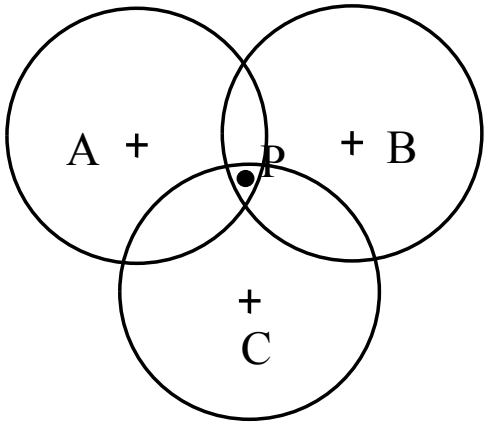


Figure 1.3: Trilateration in the presence of noisy signals.

distance measurement. This error is caused due to noise sources including atmospheric distortions such as ionospheric and tropospheric delays, relativistic effects, timing jitter, clock drift, interference and multipath [62]. In such a case the circles do not intersect at a single point as shown in Figure 1.3 and the position of the receiver is calculated by least squares fit or a similar method. In the following paragraphs, we briefly describe the noise sources present in a GNSS.

As mentioned before, the satellites have precise atomic clocks which are used to provide timing information to the receiver. Even with highly precise atomic clocks, the timing difference between the satellite and ground control stations can be as high as 1 ms. This roughly translates to a distance error of 300 km (distance = speed $\times$ time, where the speed is 3×10^5 km/s). This noise can be reduced by sending clock correction signals to the satellites periodically, typically reducing this error to the range of 1-4 m. Error is also added to the range measurements due to the difference between the expected and actual orbital position of a GNSS satellite, known as ephemeris error. This noise is small and causes an error on the order of 0.8m. The satellite clock is also affected by both special and general relativistic errors. These errors are caused by the changes in the satellite speed relative to the Earth

Noise	Error (m)
Clock	1-4
Ephemeris	1-5
Receiver	0-10
Ionosphere	10-30
Troposphere	10-30
Multipath	1-3

Table 1.2: GNSS error sources.

and in the gravitational potential. The errors due to these can reach a maximum of 70 ns (approximately 21 m in accuracy).

As the signal propagates through the troposphere and ionosphere it experiences refraction due to a change in the refractive indices of these layers. This causes a change in the speed of the signal in these layers as compared to free space. The errors caused by the atmospheric effects for each of these layers are in the range of 10-30 m.

GNSS signals can receive interference from existing narrowband signals such as amplitude modulated signals, frequency modulated harmonics, intentional continuous wave jammers, and wideband signals such as television transmitter harmonics and matched bandwidth jammers [71]. In a highly cluttered environment such as a city, the signal can bounce off (reflect from) buildings and sometimes even be diffracted. Since this signal is delayed before reaching the receiver, there is a bias in the distance measured. The errors due to interference and multipath are typically in the range of 1-3 m. In addition, the receiver may be affected by noise due to weather conditions, thermal noise jitter and inaccuracies in the antenna design. The errors caused by all these typical noise sources are summarized in table 1.1 [71]. In a modern GNSS receiver, these noises are modeled and filtered to improve the accuracy of the calculated position [71].

Several augmentations, or other sources of information, have also been combined with GNSS to improve its accuracy. For example, a differential GNSS (DGNSS)

uses an Earthbound (nearby to the receiver) tower that has a more expensive and accurate clock than the receiver to obtain an additional distance measurement [37, 59, 73, 111, 161]. It is assumed that the tower's position has been carefully surveyed and is known more precisely than its GNSS estimate. Together, these sources of information can be used to improve the accuracy of the trilateration. In GNSS-denied environments such as mountainous regions and cities, additional hardware such as an inertial measurement unit (IMU) [123], gyroscope, radio frequency (RF) signal from mobile towers [76] or wireless routers [138] can be used to provide tracking measurements. Road maps have been incorporated in modern automobile navigation systems to improve the accuracy. In this case, any position calculated by pure GNSS is pulled towards the nearest road through the use of a filtering method [48, 54]. All these techniques have led to an increase in the accuracy of GNSS tracking, thus increasing its potential applications.

1.2 Local positioning systems

An LPS as the name suggests is concerned with operations in local areas such as buildings. The GNSSs do not transmit signals with enough power to penetrate indoors. Even if the GNSS signals do penetrate indoors, they are degraded, providing accuracy (if at all acceptable) in the range of 1 – 100 m, which is unsuitable for most indoor applications. The problem of indoor navigation is similar to outdoor navigation but suffers from distinctive challenges such as high degree of clutter, non-line of sight (NLOS), placement of transmitters such that there is a good coverage of the area, the ease of installation, placement and calibration of the position of the transmitters, as well as their security. Privacy concerns and market awareness are some other challenges that have an impact on the widespread use of LPSs [78]. Over

the past couple of decades a number of technologies have been used for navigation and surveillance indoors. These include optical, pressure, IR, ultrasound, RFID, and WLAN. The following paragraphs briefly describe some of these sensors to provide an idea of their strengths and limitations in LPS applications. Some of these sensors have been discussed in greater detail by Ganjali in his thesis [45]. Mautz [102] provides a detailed survey of the range and accuracies of existing LPSs. Kolodziej and Hjelm [78] also describe in detail various sensors used in current indoor positioning systems.

Optical sensors such as cameras or photodiodes sense the amount of light they receive from the environment and convert it into voltage levels, which are then stored as different pixel values. These sensors can be used to identify the occupants of a room without their active cooperation. They can also provide information about the state of the person, such as sitting or standing. The advantage of these sensors is that they do not require the tracked person to carry an additional device. However, they require LOS signals. Occlusions can degrade performance. These sensor systems also require a significant amount of computations to process information about the color, depth and pose so as to uniquely identify occupants in the room. Microsoft's EasyLiving [15] is an example where computer vision is used to process the videos acquired by cameras to develop an indoor tracking system.

Systems such as Smart Floor developed at Georgia Tech use pressure sensors to identify and locate a person using the force signature of the person [116]. Although this system can identify and locate a person with greater than 90% accuracy, it requires placement of a large number of sensors which increases the cost.

IR sensor based positioning systems consist of an array of transmitters and receivers which are used to determine the position of the object. These transmitters emit uniquely modulated IR beacons which are used to identify the object. These beacons are not visible to the human eye. Although these devices can provide high

accuracy tracking, their range is limited to a few meters. Moreover, these signals cannot pass through opaque objects and hence requires placement of a large number of sensors leading to an increase in cost.

Ultrasound sensor systems are used in indoor ranging systems to improve the accuracy of tracking, since they can provide high time resolution. They are usually used in conjunction with other existing systems such as RF. Examples where ultrasound sensors have been used are the Active Bat [57] and Cricket [122] systems. These sensors can achieve sub-decimeter accuracy for a large number of measurements, but they face some of the same disadvantages as IR sensors, such as LOS and cost.

RFID based systems contain two basic components – scanners and tags. There are two types of RFID sensors, passive and active. Passive RFID tags are used only in conjunction with the scanners that read them. These tags do not use any batteries and they modulate the reflected signals in such a way so as to uniquely identify the tag. Passive devices have a range between 1-2 meters. Active RFID tags contain batteries to power small tranceivers that broadcast their ID which in turn is picked up by receivers. The range of active RFID can be as large as tens of meters. However, these devices can only report the presence of the tag in the area and not its exact location. Therefore, these systems are not suitable for high accuracy tracking [93]. LANDMARC is an example where active RFID has been used to determine the rough locations of objects indoors [114].

In a WLAN-based tracking system, the RF signal strength at a location is reported to a server. The location of the receiver is then estimated using RF fingerprints previously stored in a database. These systems require an extensive calibration process to develop these fingerprints [63, 70, 141]. WLAN systems use IEEE 802.11 as the networking protocol. The advantage of this system is that a large part of the backbone is already in place. An example of a tracking system using WLAN has

been developed by Ekahau [33]. This system utilizes the existing WLAN access points installed in the facility and the network cards already installed on user devices. Although these systems have low infrastructure costs, the accuracy of a typical WLAN tracking system is low (3 to 30 meters).

1.3 Ultra-wideband

The term "UWB" originated with the Defense Advanced Research Projects Agency's (DARPA) report in 1990 on the assessment of short-pulse wave technology and served as a way to distinguish it from other conventional radar technologies [44]. UWB technology has been used in indoor positioning systems since the U.S. Federal Communications Commission authorized limited use of UWB devices in 2002 [40]. UWB signals are defined as signals with ultra short pulses (< 1 ns) with a low duty cycle ($< 0.5\%$) capable of transmitting signals over a wide range of frequencies (3.1 to 10.6 GHz simultaneously) and having a large bandwidth (> 500 MHz) [93]. The UWB frequency range is shown in Figure 1.4 [82]. Due to its large bandwidth and high data rate UWB can be used for a variety of applications such as data communications for high bandwidth video and data transmission [30, 41, 121], radar for through wall imaging to detect people and objects [1, 19, 83, 110] medical imaging [11, 20, 27, 84], and localization for inventory and target tracking [2, 42, 133, 159]. Our work makes use of an UWB-based LPS. This technology has shown promise towards meeting the challenges of indoor position tracking, including ease of installation, handling NLOS and potential accuracy in tracking targets [41].

The pulses are short so that they will not disrupt other signals in the broad frequency range and avoid pulse-on-pulse overlap from neighboring transmitters. The low duty cycles enable low power consumption by UWB transmitters as compared to

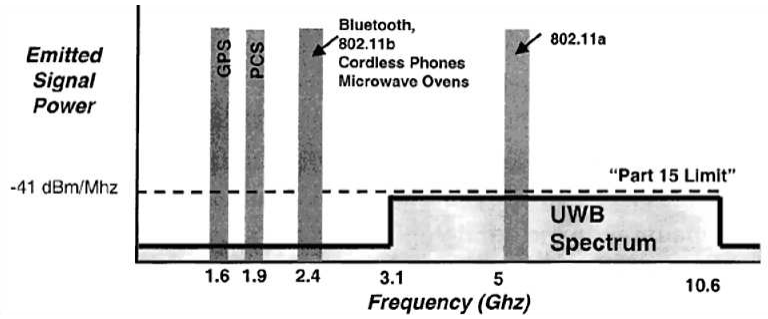


Figure 1.4: Frequency range for FCC approved UWB signals.

existing RF transmitters, making them suitable for battery operated equipment [47]. Due to the short duration of the UWB pulses, the probability of signal reflection overlap caused by multipath is reduced, making it easier to filter the original signal [49]. This property also makes UWB robust to multi-path fading as compared to other narrowband signals [50, 115, 128, 163]. The UWB signals have a very low power spectral density (PSD) since their energy is spread over a wide range of frequencies. As the PSD is below the noise threshold of existing narrowband receivers, it appears as noise to these systems. Another advantage of UWB compared to other narrowband transceivers is that UWB transmitters do not require power amplifiers or complex mixers due to carrierless transmission.

However, UWB technology is not immune to multipath and interference from other signals. UWB signals can degrade significantly when propagating in harsh environments where dense multipath is expected. For example in a shipyard where cargo containers are being tracked, UWB signals can experience a significant amount of multipath from signals reflected from the metallic walls of the containers. In addition, they can experience interference from narrowband emitters dependending on the type of antennas being used and their orientation [105]. When propagating through walls, UWB signals can suffer distortion due to the dispersive properties of the walls in addition to some attenuation [109]. The different spectral components of

a UWB signal also suffer varying amounts of delay and attenuation since the dielectric properties of the material through which they propagate behave differently at different frequencies. Due to the short, nanosecond pulses, time synchronization and sampling provides a challenge in the development of high frequency analog-to-digital converters. This makes UWB systems sensitive to timing jitter and drift.

1.4 Data fusion

The American Joint Directors of Laboratories (JDL) Data Fusion Panel [155] has defined data fusion as "a process dealing with the association, correlation, and combination of data and information from single or multiple sources to achieve refined position and identity estimates, and complete and timely assessments of situations and threats as well as their significance". Other researchers involved in multi-sensor data fusion have also defined it similarly [97,148,150]; as a tool for synergistic combination of signals from different sources to provide a better understanding of the phenomenon under consideration and to improve the quality of information.

Multi-sensor data fusion is needed when no single sensor can provide a complete picture of the phenomenon taking place. Some of the issues caused by the use of a single sensor that are resolved by use of data fusion are overcoming occlusions, reducing uncertainty in measurements, covering large tracking areas and improving precision [55, 56, 124, 148]. These issues are described briefly in the following paragraphs.

A single sensor has only a partial view of the environment and therefore may not be able to identify multiple targets in an area [18]. A complete view of the environment can be achieved by using multiple sensors. Hence multiple sensors placed at different locations can help to segment multiple targets effectively. For example,

a single camera can identify multiple targets in a room if they are well separated. However, if multiple targets are close to each other or if they are partially occluded the camera will not be able to segment them effectively [10, 66, 95]. If an additional camera is placed at a different location then the two cameras will be able to provide a better view of the environment. This method of data fusion is known as *cooperative fusion*.

A single type of sensor can provide data only in the visible range of the electromagnetic spectrum. For example a camera can only sense in the visible spectrum. If there is a change in environmental conditions, for example, poor lighting condition caused by an unexpected power failure, this sensor may not be able to operate satisfactorily. This will lead to uncertainty in measurements. Use of a different type of sensor, such as an infrared camera which can provide data in another modality can be used to reduce the uncertainty in measurements received by the first sensor type [7, 79, 143]. Other applications where multiple sensors are used to reduce uncertainty are fusion of radar with infrared for multi-target tracking [107] and fusion of synthetic aperture radar (SAR) with LiDAR for mapping forest structure for wildlife [67]. This type of data fusion is also known as *competitive fusion*.

Large-scale applications such as battlespace or traffic analysis require placement of multiple sensors in order to provide a complete coverage of the area. The sensors used in these applications may be either of the same or different types. In such applications typically a distributed architecture is used to handle the data fusion process. Hence, this type of data fusion is also known as *distributed fusion*. In such large scale applications data management issues play an important role due to the large number of sensors present in the system. Bandwidth required for communication is another issue in distributed sensor techniques [166]. Therefore the choice of the optimal sensor parameters varies significantly based on the application [157, 165].

Another type of data fusion is to improve the precision or quality of data. An example of this is D-GPS, where land based stations provide signals to correct the GPS pseudoranges [117]. The correction signals are examples of how data from multiple sensor systems (GPS and land based antennas) can be used to improve the positioning accuracy.

In the applications discussed above, data fusion occurs at different levels based on the type of information obtained by the system [31, 97, 98, 120]: *signal level, pixel level, feature level,* and *symbol level*. Signal level of data fusion can be considered as a pre-processing step where signals of different types are fused to create a new signal with a higher signal to noise ratio. For example, a combination of SAR and multispectral imaging camera can be used to provide a higher resolution image. In pixel level of data fusion, signals are fused on a pixel-by-pixel basis. It refers to the fusion of the measured physical measurements. The fusion of measurements from multi-spectral and visible range cameras is an example of pixel level fusion. At the feature level, features such as position, speed, edges, textures, etc. are extracted from the raw measurements provided by the sensors. Raw measurements from visible and infrared cameras can be processed to extract features which are used to identify the location of a target. The symbol level is where data is fused at a high level of abstraction. Data from multiple sensors are fused at this level. For example, camera data is segmented to extract the location of the target. Range information from multiple RF sensors can be combined to calculate the location of the target. The location information obtained from the camera and RF sensors are then fused to provide a better estimate of the location of the target. This level is also known as decision level.

Chapter 3 of this dissertation is concerned with data fusion combining different sensor modalities to improve tracking accuracy. In the next section we discuss sensor

ontology and introduce the distinction between aided versus unaided. We develop new methods to fuse data from aided and unaided sensors in a filtering framework.

1.5 Sensor ontology

The distinction of *passive* and *active* sensor types can be useful in the consideration of tracking problems. A passive sensor acquires data without probing the environment. An active sensor transmits and receives signals by actively probing the environment. The advantage of using passive sensors is that they can acquire data unobtrusively. Since the information being extracted by passive sensors is known only to the operator, the information is more secure. Data security can be a concern, for example in military applications where active sensing can be intercepted. Infrared (IR), ultrasound, RFID, WLAN and UWB are some examples of active sensors while optical, pressure and electromagnetic sensors are a few examples of passive sensors. Researchers have investigated combinations of active sensors [38,92,103,142,156], passive sensors [16, 24, 61, 91, 129] and hybrid active and passive sensors [34, 86, 127, 132] for target tracking [26].

Some researchers have classified sensor systems based on whether the targets are *cooperative* or *non-cooperative* [164]. A cooperative sensor system is where the target actively participates in its tracking and/or is aware that it is being tracked. In this case the target may carry an additional device. Liu et al. [94] described a non-cooperative sensor system as one where the target is passive and does not transmit signal. A non-cooperative sensor system can can also be perceived as one where the target is trying to prevent its identification on purpose. This may be true of targets in military applications. However, there is ambiguity in this nomenclature since non-cooperative can have multiple meanings. Bosch and Lescure [13] proposed

an algorithm to determine the reflection coefficient of non-cooperative targets using time-of-flight measurements of a laser rangefinder. Kim et al. [75] designed methods to measure six-degree-of-freedom displacement using cooperative targets. Gavan [46] developed a RFID system for detection and identification of remote cooperative targets.

In this work we propose a new ontology: *aided* versus *unaided*, based on whether the target aids in its tracking. An aided sensor is one that receives information from the target(s) being tracked specifically intended to assist with determining location and/or identity. An unaided sensor is one that does not receive such information. The distinction does not precisely align with cooperation. An aided sensor may be said to be tracking cooperative targets, but an unaided sensor may be tracking targets that are cooperative, uncooperative, or indifferent to the fact they are being tracked.

Figure 1.5 shows a possible sensor ontology that combines the passive/active distinction with the aided/unaided distinction, along with examples of each type.

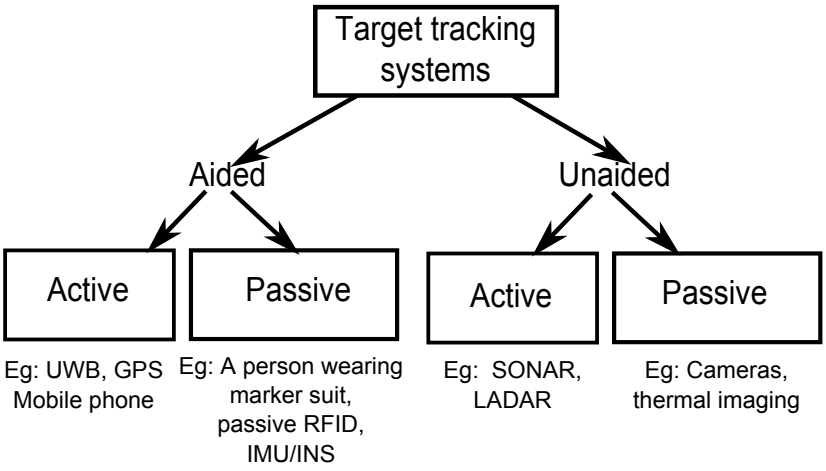


Figure 1.5: Classification of target tracking systems.

GPS, mobile phones, UWB are some examples where the target uses an addi-

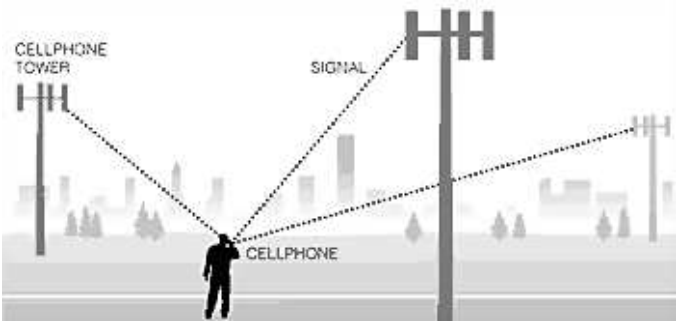


Figure 1.6: Example of aided target tracking using an active sensor.

Figure 1.7: Example of aided target tracking using a passive sensor.

tional device to broadcast information about its location. The mobile phones broadcast their unique ID to the cell towers and their position is calculated using multilateration. Figure 1.6 shows an example of a person using a mobile phone to aid tracking by broadcasting their identity (source: US District Court, Southern District of TX). Since the phone actively probes an environment using radio waves to transmit and receive data it is an example of an aided system employing active sensors.

A target wearing a motion capture/marker suit in an indoor environment whose motion is captured using a camera can be considered an example of aided target tracking system that employs a passive sensor [108]. Figure 1.7 shows a person wearing a motion capture suit [23]. The motion capture suit shown in the figure contains LEDs (Light Emitting Diodes) placed at specific locations on the suit which

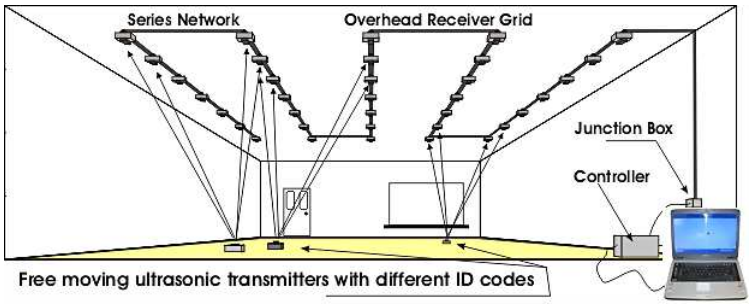


Figure 1.8: Example of unaided target tracking using active sensor.

makes it easier for a passive sensor to segment the location of the target. IMU/INS and passive RFIDs are other examples of aided tracking where passive sensors are used.

SONAR, LADAR and ultrasound are examples of unaided tracking systems where active sensors are used. In these systems the positions of the targets are detected by calculating the time between the transmission and reception of signals. The targets do not provide any information about their position. Figure 1.8 shows an example of tracking people and objects using ultrasonic sensors [60]. Here the sensors actively probe the environment to determine the positions of targets.

Figure 1.9 shows an example of unaided multi-target tracking in a room using multiple video cameras [125]. The figure shows four targets present in the room. The targets do not provide any information about their position. When they are far away from each other and do not overlap, they are easier to segment uniquely over time. If the targets overlap then it becomes difficult to distinguish them and sometimes leads to switching of target identities. In such a situation use of a different type of sensor that provides additional information can help to classify the targets.

In indoor position tracking we are concerned with solving two problems: improving precision and data association. Data association is a major problem in unaided tracking systems since the sensor does not have knowledge of the position of the

Figure 1.9: Example of unaided target tracking using a passive sensor.

target(s). The advantage of aided tracking systems is that there is no data association problem since the target broadcasts its identity. In our work we combine an aided and unaided target tracking systems to solve the data association problem and improve the precision of tracking. When the unaided system provides precise measurements as compared to the aided system, the fused output improves the precision of the aided system. On the contrary when the unaided system provides imprecise measurements the aided system is used to improve the precision of the fused tracking.

1.6 Filtering

A filter is a mathematical tool that uses an expected dynamic model to help mitigate noise in sensor data. A filter can also be used to fuse data from multiple sensors. The Kalman and particle filters are two popular filtering algorithms that are used in tracking applications. We use the Kalman filter when the system is linear and the noises are Gaussian, and the particle filter when the system is nonlinear or intractable or when the noises are non-Gaussian. The following sections describe the

filter model we use for our experiments, and provide background on the Kalman and particle filters.

1.6.1 Filter model

In the context of filtering the parameters of the model which describe the behavior of the object are known as state variables. The values of interest are assumed to lie on a distribution, rather than at a specific value. This allows for the representation of uncertainty about the values. In the case of tracking people in the X-Y plane we can consider the state variables to be x_t, y_t as the positions and $\dot{x}_t, \dot{y}_t$ as the velocities along the x and y axes at time t. The state of the system $\mathbf{X_t}$ can be written in the matrix form as

$$\mathbf{X_t} = \begin{bmatrix} x_t \\ \dot{x}_t \\ y_t \\ \dot{y}_t \end{bmatrix} \tag{1.1}$$

The state transition equations are a set of equations that describe the range of expected behaviors of the thing being tracked. Like the state variables, these equations can be anything. For a discrete 2D linear, constant velocity model, the state transition equations $\boldsymbol{f}$ can be written as [4–6, 14, 32, 74, 88, 89, 134]

$$\boldsymbol{f} = \begin{bmatrix} x_{t+1} = x_t + (\delta t)\dot{x}_t + (\delta t)u_{x,t} \\ \dot{x}_{t+1} = \dot{x}_t + u_{x,t} \\ y_{t+1} = y_t + (\delta t)\dot{y}_t + (\delta t)u_{y,t} \\ \dot{y}_{t+1} = \dot{y}_t + u_{y,t} \end{bmatrix} \tag{1.2}$$

The state transition equations use dynamic noise to describe the variability in the

possible range of next states. The state transition equations can be expressed in matrix form as

$$\boldsymbol{f} = \left[\mathbf{X}_{t+1} = \mathbf{\Phi} \mathbf{X}_t + \mathbf{B} \mathbf{U}_t \right] \tag{1.3}$$

where the state transition matrix $\mathbf{\Phi}$ can be written as

$$\mathbf{\Phi} = \begin{bmatrix} 1 & \delta t & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \delta t \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{1.4}$$

and δt is the time interval between two consecutive measurements.

The matrix $\mathbf{B}$ is written as

$$\mathbf{B} = \begin{bmatrix} \delta t & 0 \\ 1 & 0 \\ 0 & \delta t \\ 0 & 1 \end{bmatrix} \tag{1.5}$$

The dynamic noise $\mathbf{U_t}$ denotes the dynamic noise during a state transition

$$\mathbf{U}_t = \begin{bmatrix} u_{x,t} \\ u_{y,t} \end{bmatrix} \tag{1.6}$$

The dynamic noise models a potential change in velocity during each time step. The discrete, dynamic noise covariance matrix, $\mathbf{Q}(t_i, t)$ for an interval $\delta t =$

$(t_i - t_{i-1})$ is given by [32, 39, 74]

$$\mathbf{Q_t} = \begin{bmatrix} \delta t^3/3 & \delta t^2/2 & 0 & 0 \\ \delta t^2/2 & \delta t & 0 & 0 \\ 0 & 0 & \delta t^3/3 & \delta t^2/2 \\ 0 & 0 & \delta t^2/2 & \delta t \end{bmatrix} \mathbf{Q} \tag{1.7}$$

and

$$\mathbf{Q} = \begin{bmatrix} \sigma_{d_x}^2 & 0 & 0 & 0 \\ 0 & \sigma_{d_x}^2 & 0 & 0 \\ 0 & 0 & \sigma_{d_y}^2 & 0 \\ 0 & 0 & 0 & \sigma_{d_y}^2 \end{bmatrix} \tag{1.8}$$

In the covariance matrix $\mathbf{Q}$ it is assumed that there is no correlation between the x and y axes. Therefore the value of covariance between x and y axes is zero.

Sensors provide observations (measurements) of the object being tracked. The observation equations are a set of equations that describe the expected range of observations given the current state of the thing being tracking. The observation equations use observation noise to describe the potential corruption of sensed observations. At each time instant t, the set of observed values $\mathbf{Z}$ from a sensor can be written as

$$\mathbf{Z} = \begin{bmatrix} \tilde{x}_t \\ \tilde{y}_t \end{bmatrix} \tag{1.9}$$

where $\tilde{x}_t, \tilde{y}_t$ is the observed position of the object. The observation equations $\boldsymbol{g}$ are

$$\boldsymbol{g} = \begin{bmatrix} \tilde{x}_t = x_t + v_{x,t} \\ \tilde{y}_t = y_t + v_{y,t} \end{bmatrix} \tag{1.10}$$

where v_x and v_y describe the expected observation noise of the sensor. The observation equations can be written in the matrix form as

$$\boldsymbol{g} = \begin{bmatrix} \mathbf{Z}_t = \mathbf{O}\mathbf{X}_t + \mathbf{N}_t \end{bmatrix} \tag{1.11}$$

where the observation matrix is given by

$$\mathbf{O} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{1.12}$$

The observation noise matrix $\mathbf{N}$ is written as

$$\mathbf{N} = \begin{bmatrix} v_{x,t} \\ v_{y,t} \end{bmatrix} \tag{1.13}$$

and the measurement noise covariance matrix $\mathbf{R}$ is given by

$$\mathbf{R} = \begin{bmatrix} \sigma_{n_x}^2 & \sigma_{n_x,n_y} \\ \sigma_{n_x,n_y} & \sigma_{n_y}^2 \end{bmatrix} \tag{1.14}$$

where $\sigma_{n_x}^2$ and $\sigma_{n_y}^2$ are variances and σ_{n_x,n_y} is the covariance along the X and Y axes respectively.

1.6.2 Kalman Filter

In a problem where the state transition and observation equations are linear and the dynamic and observation noises are Gaussian, the Kalman filter provides an ideal estimate of the state of the object under consideration. The Kalman filter provides a filtered output by weighting measurements against predictions based upon measurement and dynamic noise covariances. In order to start the Kalman filter we need some initial position for the state $\mathbf{X}$. We can assume that the initial state is provided by the problem definition or the first observation when tracking begins. Hence intial state can be expressed as

$$\mathbf{X} = \begin{bmatrix} x_0 \\ \dot{x}_0 \\ y_0 \\ \dot{y}_0 \end{bmatrix} \tag{1.15}$$

where x_0, y_0 is the initial location of the person along the X and Y axes. In our example, we consider the initial velocities to be zero. The initial prediction is the same as the initial state.

The first step in the operation of Kalman filter is the calculation of the Kalman gain matrix. The gain matrix gives the weighting between measurements and estimates which can be calculated as

$$\mathbf{K}_t = \mathbf{S}_{t,t,-1}\mathbf{O}^T[\mathbf{O}\mathbf{S}_{t,t,-1}\mathbf{O}^T + \mathbf{R}]^{-1} \tag{1.16}$$

where, $\mathbf{K}$ is the Kalman gain and $\mathbf{S}_{t,t-1}$ is the predictor error covariance at time t and $\mathbf{O}$ is the observation matrix.

The next step is updating the state and predictor covariance matrix. The state is updated using the following equation

$$\mathbf{X}_{t,t} = \mathbf{X}_{t,t-1} + \mathbf{K}_t[\mathbf{Z}_t - \mathbf{O}\mathbf{X}_{t,t-1}] \tag{1.17}$$

and the predictor covariance matrix is updated using the following equation

$$\mathbf{S}_{t,t} = [\mathbf{I} - \mathbf{K}_t\mathbf{O}]\mathbf{S}_{t,t-1} \tag{1.18}$$

The final step consists of predicting the next state and predictor covariance matrix given by

$$\mathbf{X}_{t+1,t} = \mathbf{\Phi}\mathbf{X}_{t,t} \tag{1.19}$$

where $\mathbf{\Phi}$ is the state transition matrix and T is the time interval between two consecutive measurements.

Finally, the predictor covariance matrix is predicted as follows

$$\mathbf{S}_{t+1,t} = \mathbf{\Phi}\mathbf{S}_{t,t}\mathbf{\Phi}^T + \mathbf{Q}_t \tag{1.20}$$

The basic Kalman filter is suitable only for applications where the transition and observation equations are linear and the noises are Gaussian. The EKF allows for non-linear equations by calculating Jacobians at each time step to linearize the problem [153]. The unscented transform [147,151] is an improvement on the EKF. It calculates sigma points from the state and covariance matrices by transforming them through the state transition equation and then rebuilding the state and covariance matrices. However, all these methods break down when the equations are intractable or when the distributions are non-Gaussian.

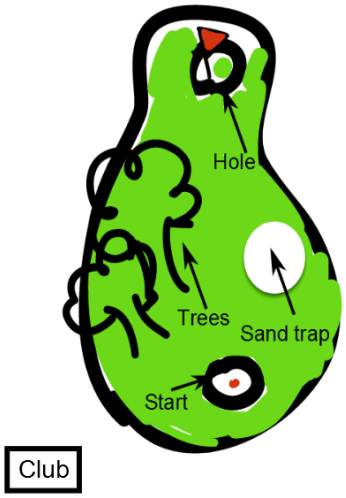


Figure 1.10: Layout of a golf course.

1.6.3 Particle Filter

Particle filtering is a sequential Monte Carlo methodology where the posterior density function is recursively approximated using a set of random samples and weights and the estimates are computed based on these samples and weights [3, 29]. The advantage with this approach is that it can use any distribution function even when the mathematical formula for the distribution function is intractable [64]. By intractable we mean that the distribution cannot be modeled easily using an analytic function. The probability distribution is approximated using a set of particles. Each particle has a state and a weight. The summation of all the particles provides an approximation to the distribution.

The concept can be explained by the following example. Consider a golf course where a player hits a ball along a course as shown in figure 1.10. Assume that it is a par-30 hole ($\approx$ 2500 m in distance; a typical par-3 hole is 230 – 270 m long). Here, the state is the location and velocity of the ball along x and y axes can be assumed

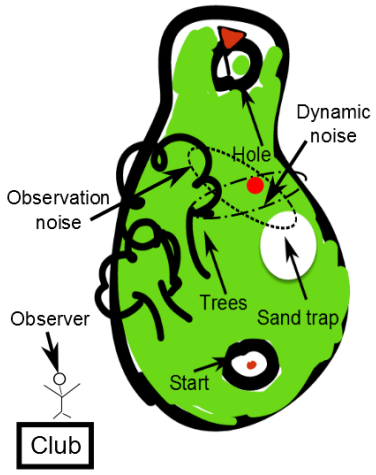


Figure 1.11: Illustration of observation and dynamic noise pdfs while playing golf.

to be similar to that given by equation 1.1.

We know that there will be variability in every shot that the player makes due to changes in motion, wind conditions and other factors. These factors contribute to the dynamic noise in the system. Let's assume that the state transition equations $\boldsymbol{f}$ are the same as in equation 1.2.

We seek to understand the distribution of potential shots. To do this, the player is asked to hit 10 balls from the same location. We consider these balls to be the set of particles used to approximate where the ball will land. As the number of shots is increased, we get a better approximation of where the ball is expected to land. The shots initially have the same probability (weight) where they could land (0.1 in this case, since the player was given 10 attempts). Assuming the player to be good at the game, the probabilities of the ball ending up on the "green" are higher than in the sand trap or the trees.

Suppose that the player is now playing a game with his friends and we would like to predict the location of the ball. We are observing the player's shots from the

club using binoculars. This can be seen in Figure 1.11. Our observation is partially occluded by the trees on the course, the distance of the ball from the club and other factors. These factors contribute to noise (errors) in the observation. Figure 1.11 shows the observation and dynamic noise distributions and the actual location of the ball. The actual location of the ball is shown by a red circle.

Now we need to answer the question, given the current location of the ball, what's the next probable location of the ball? We start with some initial position (here, the 'tee-off' location) and then depending on the observed location of where the ball was hit in each round and previous knowledge of the player's shot potential (dynamics) we update our next estimates. Using the recursive Bayesian estimation we can write

$$p(x_{t+1}|z_{t+1}) = \frac{p(x_{t+1}|x_t)p(z_{t+1}|x_{t+1})}{p(z_{t+1}|z_t)}p(x_t|z_t) \tag{1.21}$$

where, x_t is the current state, z_t is the current measurement and x_{t+1} is the next state.

For the next round of shots, we would give higher weight to the shots where the balls ended up on the green than to the shots where the balls ended up in the sand trap or the trees due to the expected (previously modeled) dynamic noise, balancing it against the known observation noise and the actual observation. Since the transition from $p(x_t|z_t)$ to $p(x_{t+1}|z_{t+1})$ is often intractable we use a set of particles to approximate the distribution. Depending on how closely these particles transition compared to the actual transition of the system, they receive higher weights.

Mathematically, we can denote the set of particles as,

$$\chi = \left\{\mathbf{X}^k, w^k\right\}_{k=1}^{K} \tag{1.22}$$

where, K is the number of particles, $\mathbf{X}^k$ is the state of each particle, and w^k is the weight of each particle.

Now, the expected value of the state can be approximated by using χ as

$$E[\chi] = \sum_{k=1}^{K} \mathbf{X}^k w^k \tag{1.23}$$

Note that the expected value is assumed to be the weighted mean of the state. This may not be the only desired answer; for example we can use the highest weighted particle to be the desired answer.

Now the problem is how to select particles from an intractable distribution $p(x|z)$ and sample from it? The answer is to introduce a known, easy-to-sample, proposal distribution $q(x|z)$. The most common approach to selecting $q()$ is called importance sampling [80] which reduces the weight update equation to

$$w_t(x_t) = w_{t-1}(x_{t-1})p(z_t|x_t) \tag{1.24}$$

In order to implement the particle filter we start by choosing the number of particles used to approximate the probability distribution function (pdf). It is chosen to be a sufficiently large number, say N=1000. The computation time increases with an increase in the number of particles. Hence, the number of particles should be chosen such that there are sufficient particles to approximate the pdf and the computation time is not extremely long.

In the absence of any knowledge about the initial state of the particles they can be initialized to the first observation, with equal weights.

$$\chi = \left\{\mathbf{X}^k, w^k\right\}_{k=1}^{N} = \left\{0, \frac{1}{N}\right\} \tag{1.25}$$

The first step at each time instant is to transition each particle with a different dynamic noise

$$\left\{\mathbf{X}_t^k\right\}_{k=1}^N = f\left\{\mathbf{X}_{t-1}^k, u_t^k\right\}_{k=1}^N \tag{1.26}$$

where, f is the matrix of state transition equations.

The next step is to update the weights for the particles using the new observation

$$w_t^k = w_{t-1}^k p(\mathbf{Z}_t|\mathbf{X}_t^k) \tag{1.27}$$

After the weights have been computed, it is necessary to normalize the new weights so that the weights add up to 1. This gives the probability of each particle

$$\left\{w^k = \frac{w^k}{\sum_{k=1}^N w^k}\right\}_{k=1}^N \tag{1.28}$$

Finally, the desired output is computed, such as the weighted average.

$$E[\mathbf{X}] = \sum_{k=1}^N \mathbf{X}^k w^k \tag{1.29}$$

At each time step each particle undergoes a transition to a new state. The new state comes from the range of possibilities defined by the state transition equations. Those particles with a transition close to that of the actual transition corresponding to the observation receive higher weights, while those with significantly different transitions receive lower weights. After a few transitions, some particles tend towards zero weights. The net result is that after some time only a few particles contribute to the approximation of the pdf. In this situation, it is necessary to resample the particles. Resampling is a process by which particles with low probabilities are "killed" and particles with higher probabilities are split up into newer particles to provide a

better approximation of the pdf. Various techniques are used to resample the particles [126]. We use one of the most common techniques called *select with replacement* in our experiments [3].

1.7 Previous Work

We use noise models and augmentations in a filtering framework to implement our ideas for improving UWB based tracking. In the following sections we describe previous work conducted by researchers in UWB based positioning systems to model the sources of noise and augment them using other sensors.

1.7.1 Noise Modeling

Suski et al. [140] enumerated six sources of noise prevalent in UWB-based indoor tracking systems: NLOS, multipath, synchronization, antenna effects, peak detection and sensor placement. NLOS errors lead to an increase in the time-of-flight measurements between tags and sensors which can bias range measurements. Multipath from fixed infrastructure such as stone walls, metallic railings (which tend to attenuate high frequency pulses) can cause similarly increased time-of-flight measurements. Synchronization is important to TOA and TDOA systems as they depend on accurate clock references to determine range. A UWB system operates over very short distances as compared to a GNSS, making the timing information even more sensitive to jitter and drift. The design and placement of antennas has an affect on the received power of signals. Accurate peak detection is also affected by TOA and TDOA measurements.

Previous works that studied the effect of noise sources on the accuracy of UWB indoor position tracking can be sub-divided by the approach taken: development of

prototypes and simulations. In the former case researchers have focused on artificially isolating the sources of error. For example, Fontana et al. [43] implemented one of the first UWB based localization system for tracking assets. The system consisted of four receivers spread over a test area approximately 30 $\times$ 15 m in dimension. They also conducted indoor and outdoor tests to determine the maximum range over which the signals could be transmitted. It was observed that the range decreased considerably indoors due to attenuation from walls and furniture placed in the laboratory.

Guoping and Rao [53] developed a low cost localization system along with a delayed correlation detector and high-speed Analog to Digital Converter (ADC) to reduce the ranging errors caused in TDOA. They conducted experiments to track the tag along the x and y axes in a 5 $\times$ 6 m area and achieved an accuracy of less than 30 cm.

Zetik et al. [160] examined mitigating the effects of multipath, synchronization, antenna effects and peak detection in a relatively controlled environment. They built a custom designed multi-channel UWB localizer and conducted experiments that employed both active and passive approaches to measure the 2D and 3D position of a moving object. However, experiments were conducted in a small area (2m $\times$ 4m) and NLOS conditions were not considered. Meier et al. [104] also developed a prototype of UWB transceiver system which suppressed multipath effects in highly reflective laboratory environments by detecting the LOS and NLOS signals and achieved sub-millimeter accuracy by filtering the measurement noise using a Kalman filter.

Mahfouz et al. [100] conducted experiments in a small test area and concentrated on reducing the effects of multipath, synchronization, antenna effects, peak detection and sensor placement on UWB localization. In order to determine the impact of each noise source, they conducted 1D, 2D and 3D experiments to measure the position of the transmitter over a small (1 - 2 m) unobstructed area. Using precision

optical tools, ground truth positions were established with an accuracy of less than 1 mm. From their experiments on mitigating the effects of different sources of error, they approached millimeter level accuracies. Zhang et al. [162] proposed an architecture for combining traditional energy-based and carrier-based detection schemes to minimize carrier phase noise and timing error effects and achieve millimeter accuracy for static and dynamic tracking. Using this architecture, they achieved root mean square errors between 2 and 6 mm for 3D tracking. Further research in [158] proposed a modified correlation algorithm to achieve sub-millimeter positioning accuracy and reduce the computational burden of systems.

Low et al. [96] reported sub-decimeter accuracy by using a combination of peak search and match filtering in LOS signals. They conducted experiments in realistic environments such as large open areas to dense office spaces to observe the effect of multipath. However, the measurements were conducted in LOS conditions to reduce large biases in range measurements caused by NLOS.

Muqaibel et al. [109] conducted experiments on ten commonly used building materials to study the impact of these materials on the propagation of UWB signals. This is important since UWB signals suffer distortion not only in magnitude but also in phase due to the large range of frequencies over which it operates. From the experiments, they observed that each spectral component of the UWB signal undergoes a significant delay and attenuation due to the dispersive effects of the material.

Simulations have also been used to demonstrate a proof-of-concept, such as by Lie et al. [90] combined an envelope detector to a leading-edge pulse detection method using a tunnel diode to achieve centimeter level accuracy for UWB ranging in a multipath environment. Jing et al. [68] studied the theoretical impact of propagation of signals through various building materials. Shen et al. [130] proposed

a new method to identify and mitigate the effects of NLOS signal propagation in an UWB based indoor localization system by comparing the mean square errors of the range estimates with the estimated LOS ranges. They conducted simulations to compare the performance of the proposed method with other methods in this area. Caron et al. [17] proposed different particle filters which can handle synchronous and asynchronous measurements received from different sensors in a multisensor system. These particle filters can switch between different observation models and also handle cases where sensors fail or their functioning changes. Denis et al. [28] used a modified Extended Kalman Filter (EKF) and modified regularized particle filter to track biases caused by transitions from LOS to NLOS and from NLOS to NLOS environments.

1.7.2 Augmentation

A few UWB augmentation techniques have been investigated. Jourdan et al. [69] combined measurements from an IMU and UWB to correct biases due to NLOS and incorrect beacon modeling in a particle filtering framework and improve the tracking accuracy of the object. The IMU provides the attitude and accelerations which is used to determine the ranges from the fixed positions of the UWB beacons, which are then used to update the position of the agent. The IMU measurements are used to discard the outliers caused by NLOS in UWB measurements. They conducted simulations to test the effectiveness of their method to simultaneously update the ranging errors and the position of the agent.

Similar work was conducted by Pittet et al. [119] by combining micro electro mechanical sensors (MEMS) inertial data along with UWB location data to reduce localization errors caused by multipath, biases due to NLOS and non-optimal channel modeling, and outliers. They used the EKF to combine the measurements received

from the sensors and improve indoor tracking of people. Their experiments were conducted in a classroom made of steel walls and approximately 12 × 12 m in dimension. Even in such an environment where heavy multipath is expected due to reflections from the steel walls, they achieved an accuracy of less than 1 m by combining measurements from the two sensors.

Corrales et al. [25] also combined an IMU and UWB to track human operators and their actions in an environment which requires active cooperation with robots. They attached 18 IMUs along with a UWB tag on a suit which is worn by the operator. The measurements received from the IMUs and the UWB tag are combined using a Kalman filter. Using a combination of the two measurements, they reduced the errors from 56 cm to 14 cm.

Cheok et al. [21] combined measurements from UWB, wheel speed encoders and a digital compass using an EKF to improve the accuracy of tracking a mobile robot. They also used fuzzy logic to remove the outliers received from UWB sensors.

MacGougan and O'Keefe [99] achieved centimeter-level accuracy in outdoor navigation by use of UWB ranges in addition to GPS signals to correct for bias and scale factor estimation.

1.8 Contributions

This work focuses on two new approaches taken to improve the precision of an UWB based indoor positioning system for tracking people. The main contributions of this dissertation are listed as follows:

1.8.1 Set switching noise

We identify and explore the effect of noise caused by the switching of sensor sets on the precision of an UWB based tracking system. Although this noise exists in GNSSs, its effect is much more pronounced in an LPS. This is due to the scale of a LPS (meter-level) compared to that of a GNSS (kilometer-level). The speed of operation of a LPS is also higher than in a GNSS causing the sensor sets to switch at a faster rate. In chapter 2 we begin with a description of sensor set switching followed by mathematical modeling of the noise due to the switching of sensor sets. Then we describe a particle filter that can help mitigate this noise. We chose the particle filter to develop a general framework that can handle Gaussian and non-Gaussian noises, and both linear and non-linear motion models. Finally we present experimental results to demonstrate the impact of this noise and how its effect can be reduced by using a particle filter. The performance is evaluated by comparing the filtered outputs to the raw measurements for dynamic position tracking.

1.8.2 Augmentation

We develop a filtering framework to combine position data from aided and unaided sensors. The goal of the proposed framework is to improve the precision of tracking and to help solve data association issues when multiple targets are present. In chapter 3 we begin with a description of the fusion framework followed by a description of multi-object tracking metrics and methods. We present experimental results on real data by varying the number of people in a room from 1 to 8 to demonstrate the effect of fusion on multi-object data fusion and precision. The performance is evaluated by using the multi-object tracking (MOT) metrics defined by the Classification of Events, Activities and Relationships (CLEAR) workshop [22, 135, 136].

Chapter 2

Sensor set switching noise

2.1 Motivation

Trilateration-based tracking relies upon measuring the distances from a fixed set of points ("sensors") to an object of interest ("tag"). In this work we consider the noise at the level of a set of sensors used in a single trilateration calculation. This noise changes when the set changes. In a GNSS, the set of sensors changes slowly because of the scale of the tracking system (see Figure 2.1). In an unobstructed area, the set of visible satellites changes approximately every 15 minutes [62]. However, in indoor UWB indoor position tracking, sensor sets change with every new measurement (typically 100 ms). The sets change while moving around a single room, and sometimes even while standing still, depending upon the received signal strengths as shown in Figure 2.2.

The problem caused by set switching is illustrated in Figure 2.3. The sequence shows three consecutive trilateration calculations that use different sensor sets, each resulting in a different tracked location, even though the object of interest has not moved. At time t, the position of the tag is computed from sensors A, B and C.

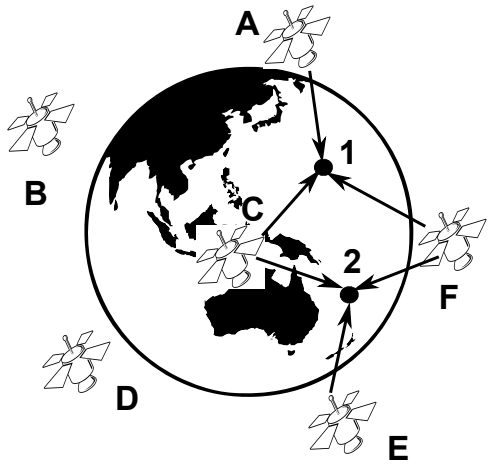


Figure 2.1: Trilateration using a GNSS (earth-sized).

At time t+1, a new sensor set consisting of A, B, C and D is used to calculate the position of the tag. It can be observed that a change in the sensor set has caused a shift in the calculated position of the tag, due to the changing collective set of noises in the distance measurements. At time t+2, a new set of sensors consisting of A, B and E causes another shift in the calculated position. Hence, switching between sensor sets at each time instant adds a different noise to the measurements corresponding to the noise model of each sensor set. This causes a "jump" in the calculated position of the tag, even when the tag is not moving. A video of such a behavior occurring at our facility using a real UWB position tracking system can be seen at http://youtu.be/B-oCDTBQLd4.

In this chapter we identify noise due to the switching of sensor sets. This noise is present in any trilateration-based tracking system, but its effect is much more pronounced in an indoor positioning system. In preliminary work our group examined this issue in simulations [45]; this work is the first to study it in a real system. After providing a mathematical model for this noise, we describe a particle filter for reducing its effect. We then demonstrate the operation of this particle filter

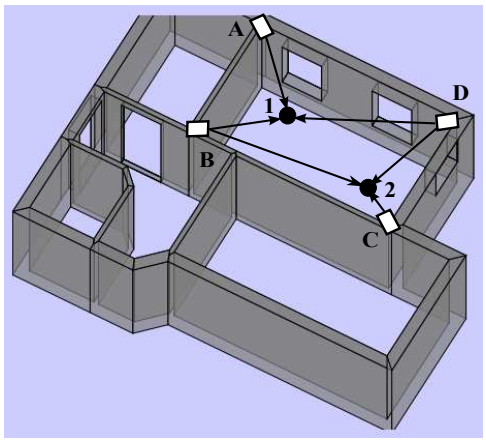


Figure 2.2: UWB indoor trilateration (building-sized).

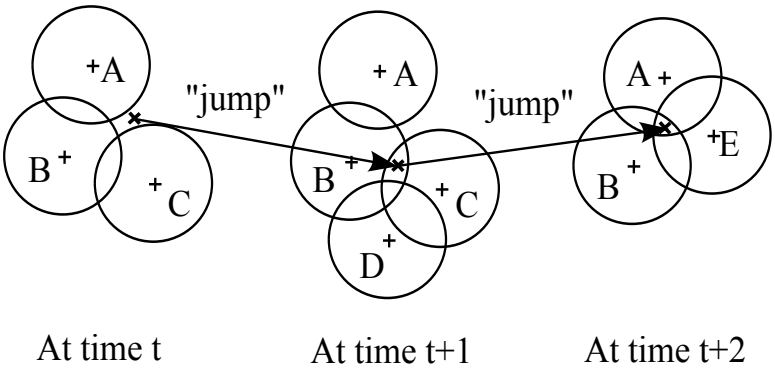


Figure 2.3: Changing noise due to sensor set switching.

on data from a real system, showing an approximately 15% improvement in accuracy over the raw measurements.

2.2 Methods

2.2.1 Facility

Our facility is located in the basement of Riggs Hall at Clemson University. The test area covers approximately 8 m × 8 m, covering the majority of a laboratory and part of an adjacent hallway. Figure 2.4 shows a picture of part of this area, where it can be observed that the laboratory and the adjacent hallway are separated by a concrete wall which is approximately 20 cm thick. The walls are approximately 5 m high with false ceilings at a height of 3 m. The false ceilings are made up of thermocol and placed on metal railings. In addition, there are two metal mailboxes and a vending machine in the hallway, and two cupboards in the laboratory. Figure 2.5 shows the locations of furniture, walls and sensors in the test area.

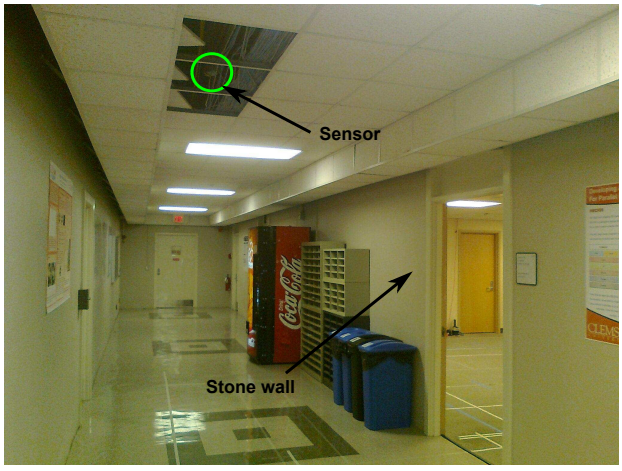


Figure 2.4: The facility and location of one of the eight sensors.

2.2.2 UWB positioning system

We used a commercially available UWB based local positioning system developed by Ubisense Inc. (Cambridge, U.K.). We installed eight Series 7000 sensors [145] in the facility at fixed locations. These sensors detect UWB pulses from Ubisense tags [144], which are tracked moving throughout the test area. Sensors are powered over network cabling using a Power-over-Ethernet switch. The Ubisense system uses a combination of angle of arrival and time difference of arrival, followed by multilateration or hyperbolic positioning, to calculate the position of a tag [146]. The master sensor chooses five sensors which have the highest demodulation power. We refer to this collection of five sensors as a sensor set.

Figure 2.5 shows the positions of the eight sensors distributed across our facility. The company recommends an install where the sensors are placed in a rectangular pattern surrounding the area of interest, with minimum NLOS conditions. However, the promise of UWB indoor position tracking is that it can be accomplished without direct LOS between the tracked object and fixed sensor points. Our install purposefully introduces some NLOS conditions from the facility in order to explore this challenge. However, it must be noted that NLOS is not the only noise source that contributes to errors in distance measurements for trilateration, and that even with a completely LOS install, we have observed significant sensor set switching noise.

The sensors are connected in a daisy-chained fashion with the master sensor providing the timing signals. These sensors are connected to a central switch which is then connected to a computer where the location engine software computes the position of the tags. Figure 2.6 shows a block diagram of the connection between the sensors and the computer which processes the positions of the tags.

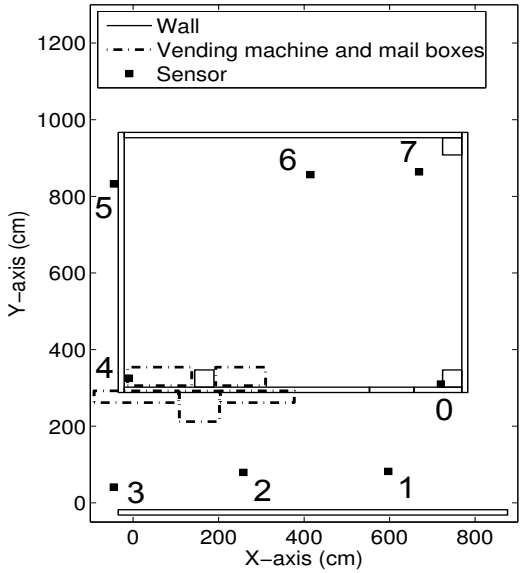


Figure 2.5: Layout of the facility (0–7 indicate UWB sensor positions).

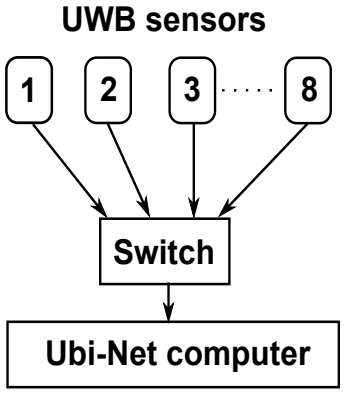


Figure 2.6: Block diagram of the configuration of the UWB system.

2.2.3 Noise model

2.2.3.1 Sensor set

We assume that a tracking measurement is calculated from a subset of the available sensors and each sensor set has a noise model associated with it. We model the noise associated with each sensor set independently. For the sake of simplicity, we assume that the noises are non-zero mean Gaussian, but our methods could be applied with other distributions.

We model the total set of I sensors as $\{1, 2, ..., i_{max}\}$. Let a sensor set s represent any subset of size ≥ 5 sensors drawn from I, denoting a specific sensor subset. We calculate a Gaussian noise model for measurements relative to their actual location for each sensor set s as

$$\mathcal{N}(\mu_x^s, \mu_y^s, \sigma_x^s, \sigma_y^s) \tag{2.1}$$

The total possible subsets can grow large as the number of sensors I grows. For example, if $\parallel I \parallel = 8$, then there are a possible total of $s_{max} = \binom{8}{5} = 56$ sensor sets. However, we assume that a relatively small number of sensor sets dominates the possibilities used for tracking measurements. Figure 2.2.3.1 shows a plot of 1,000 measurements made by our Ubisense system, all at a single ground truth location shown by a '+'. The dots represent the actual measurements received from the system. Figure 2.8 shows the frequency distribution of the sensor sets for these measurements. The most common 5 sensor sets account for 858 measurements, or 85.8% of the data. These 5 sensor sets are '76540', '75420', '65410', '65420' and '76541', where the numbers indicate the sensor (see figure 2.5) used to provide a measurement. The most commonly occurring sensor set corresponds to the sensors that are most LOS, and

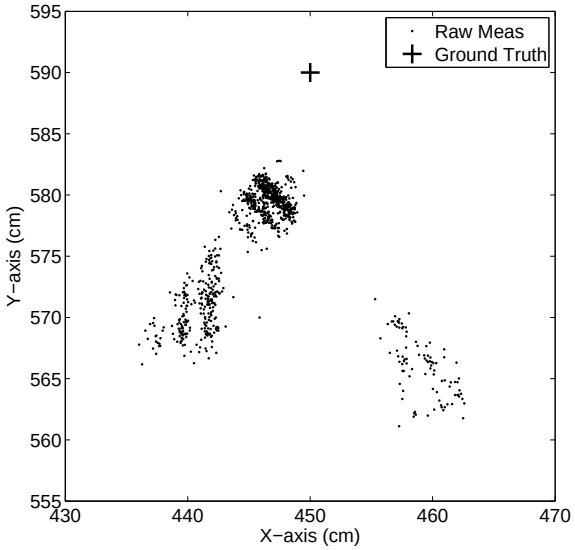


Figure 2.7: 1000 measurements collected at a single location (450, 590, 92) cm.

therefore generally the most powerful signals. However, it accounts for less than 45% of the total data. The second most commonly occurring sensor set contains a sensor which lies in the hallway and contributes approximately 20% of the measurements. Similarly, other sensor sets contain at least one sensor which lies in the hallway, providing better angular coverage but more NLOS conditions.

Figure 2.9 shows the noise models for four sensor sets from the data collected in Figure 2.2.3.1. The noise model for each sensor set is given by $(\mu_x^s, \mu_y^s, \sigma_x^s, \sigma_y^s)$ where (μ_x^s, μ_y^s) corresponds to the average shift of the measurements from the sensor set relative to the ground truth location, and (σ_x^s, σ_y^s) corresponds to the standard deviation of the measurements from their mean. The length of the axes of the ellipses in figure 2.9 correspond to three standard deviations.

2.2.3.2 Calibration

In order to calculate our noise model, we conduct a calibration step. A tag is placed at a known location, and 1000 measurements are collected. This process

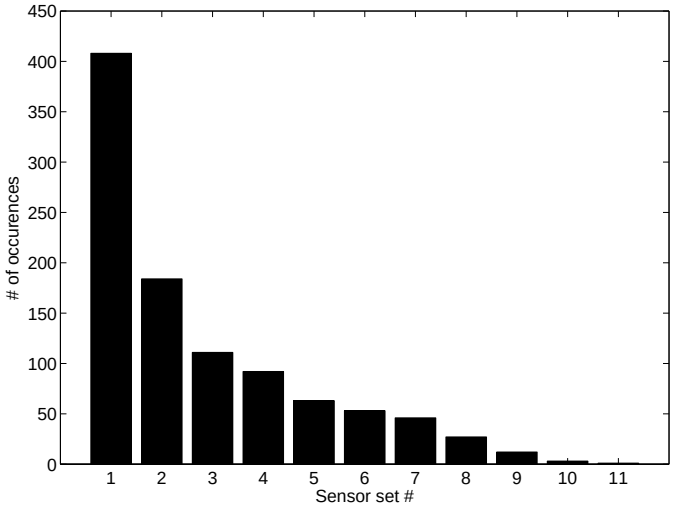
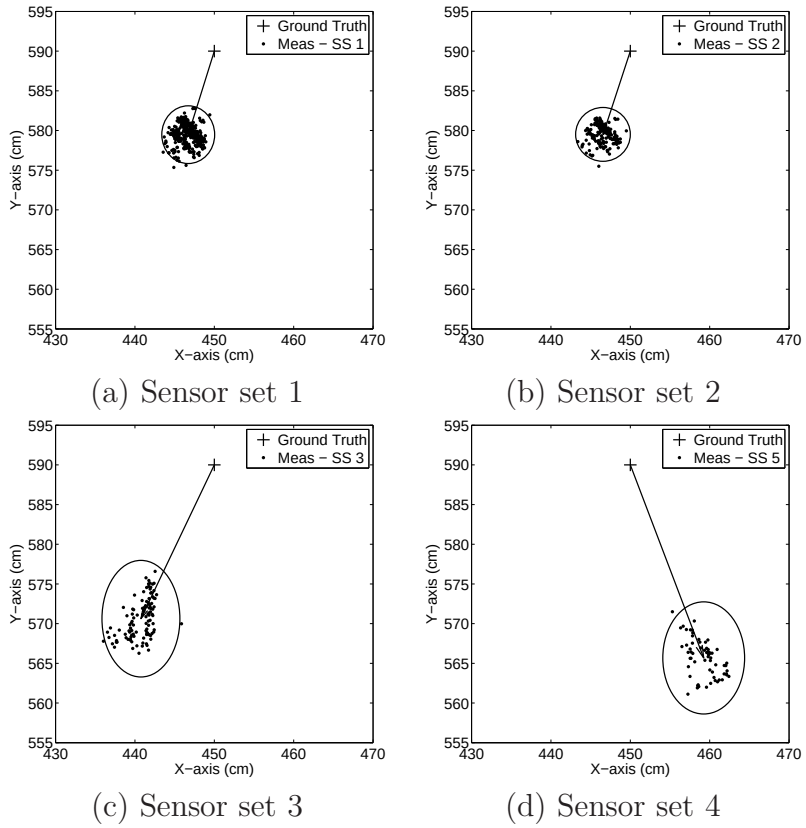


Figure 2.8: Distribution of measurements.



(a) Sensor set 1 (b) Sensor set 2

(c) Sensor set 3 (d) Sensor set 4

Figure 2.9: Noise models of four sensor sets at (X,Y,Z) = (450,590,92) cm.

is repeated at 6 different locations distributed throughout the facility. The noise parameters $(\mu_x^s, \mu_y^s, \sigma_x^s, \sigma_y^s)$ for each sensor set are calculated at each location, and then weighted-averaged by the number of measurements for each sensor set across the 6 locations. At a single location, if a sensor set has less than 30 measurements then no noise model is calculated at that location. After weighted-averaging, some sensor sets may have no model. We therefore also calculate a facility-wide noise model that is used by default for measurements taken from a non-modeled sensor set. The facility-wide noise model is taken as the average of all measurements taken during the calibration step.

Figure 2.10 shows how calibration data was collected. A tag was placed on a wooden sawhorse 92 cm in height. This height was chosen since it corresponds to the typical height of the waist of an upright person. The base of the setup is marked at intervals of 10 cm so that it can be aligned with markings on the floor of the facility. This enables accurate positioning of the sawhorse with respect to the ground truth coordinate system. We used laser levels and tape measures to ensure that the ground truth locations are accurate to within 1 cm relative to the calibrated Ubisense coordinate system.

2.2.4 Basic particle filter

A basic particle filter algorithm is described in section 1.6.3. In our experiments we compare the performance of our set noise particle filter to this basic particle

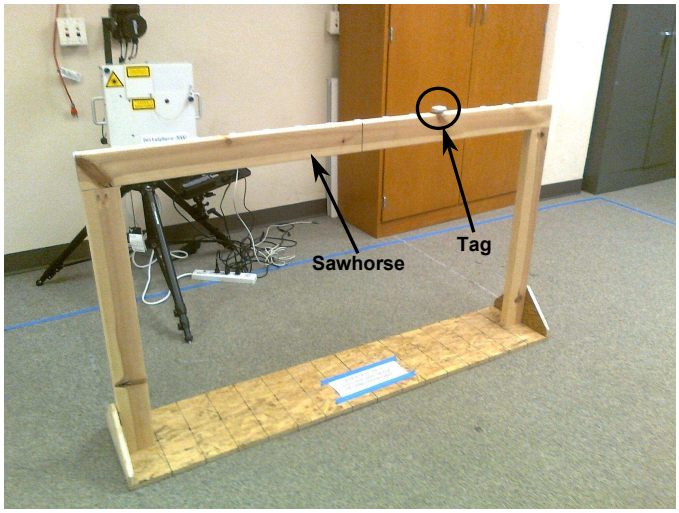


Figure 2.10: Setup used to collect calibration data.

filter. We initialize all particles to have equal weight and the same initial state:

$$\chi = \left\{\mathbf{X}^k, w^k\right\}_{k=1}^{K} = \left\{ \begin{bmatrix} x_0 \\ 0 \\ y_0 \\ 0 \end{bmatrix}, \frac{1}{K} \right\} \tag{2.2}$$

where x_0 and y_0 are the known starting position of the tag along the x and y axes with zero initial velocities and K is the number of samples; in this work we use $K = 1000$. Here the dynamic noise is chosen to be zero-mean Gaussian random variables with a standard deviation of $\sigma_{d_x}, \sigma_{d_y}$ along the X and Y directions.

2.2.5 Set noise particle filter

Using our sensor set noise model, the basic particle filter algorithm is adjusted as follows. At each time t, the set of observed values $\mathbf{Z}$ given by equation 1.9 is

modified as

$$\mathbf{Z} = \begin{bmatrix} \tilde{s}_t \\ \tilde{x}_t \\ \tilde{y}_t \end{bmatrix} \tag{2.3}$$

where $\tilde{s}_t$ is the sensor set used to measure $\tilde{x}_t, \tilde{y}_t$. The observation equations $\boldsymbol{g}$ given by equation 1.10 are rewritten as

$$\boldsymbol{g} = \begin{bmatrix} \tilde{s}_t \leftarrow \{1, 2, ..., s_{max}\} \\ \tilde{x}_t = x_t + \mathcal{N}(\mu_x^{\tilde{s}_t}, \sigma_x^{\tilde{s}_t}) \\ \tilde{y}_t = y_t + \mathcal{N}(\mu_y^{\tilde{s}_t}, \sigma_y^{\tilde{s}_t}) \end{bmatrix} \tag{2.4}$$

where it is assumed that a random non-zero mean Gaussian noise associated with sensor set $\tilde{s}_t$ has been added to the actual position to produce the measurement.

The weight update step given in equation 1.27 is replaced with

$$p(\mathbf{Z}_t|\mathbf{X}_t^k) = \exp - \left(\frac{((x_t^k - \mu_x^{\tilde{s}_t}) - \tilde{x}_t)^2}{2(\sigma_x^{\tilde{s}_t})^2} + \frac{((y_t^k - \mu_y^{\tilde{s}_t}) - \tilde{y}_t)^2}{2(\sigma_y^{\tilde{s}_t})^2} \right) \tag{2.5}$$

where $(x_t^k - \mu_x^{\tilde{s}_t})$ and $(y_t^k - \mu_y^{\tilde{s}_t})$ gives the most probable measurement of each particle. Equation 2.5 calculates the likelihood of obtaining the actual observed measurement relative to the most probable state of the particle, according to the measurement noise distribution associated with the sensor set used to take the measurement.

All the other steps are the same as described for the basic particle filter.

2.2.6 Data collection

Figure 2.11 shows the apparatus used to record experimental data for testing. A tag was placed on a tripod resting on a trolley. The tripod was adjusted so that it would match up to the same height (92 cm) used to collect calibration data. The

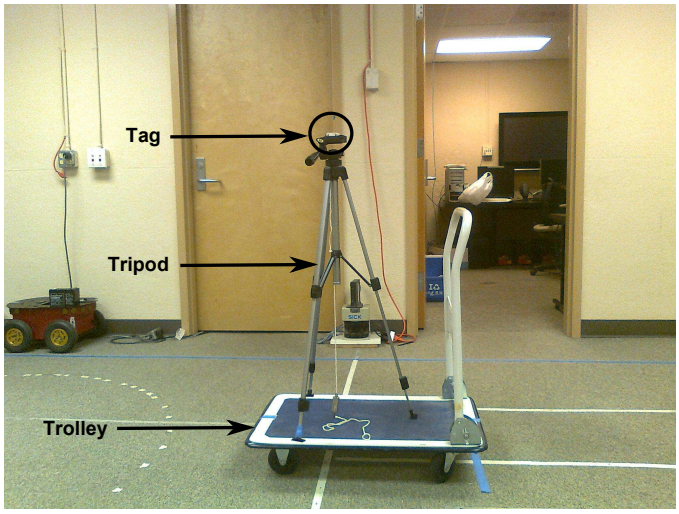


Figure 2.11: Setup used to collect recordings.

trolley was then pulled manually along a track laid on the ground at different speeds. For each recording, the apparatus was pulled back and forth seven times along a 250 cm straight line. The total distance covered in each recording is 1750 cm.

Figure 2.12 shows the location of the track in the test area. For each recording along the track, the Ubisense system provides raw measurements of the tag along the X and Y axes, and the sensors used to calculate each measurement. We turned off the simple averaging filters provided by the Ubisense system and collected the raw measurements.

Five recordings were collected along the track at different speeds. The speed was varied from ≈11 cm/s (extremely slow motion) to ≈120 cm/s (walk speed) [77, 152]. Table 2.1 lists the recording number and the approximate speed of the recording. The speeds were chosen to test the viability of our method for a range of motion dynamics resembling a slow moving robot to the walking of a person.

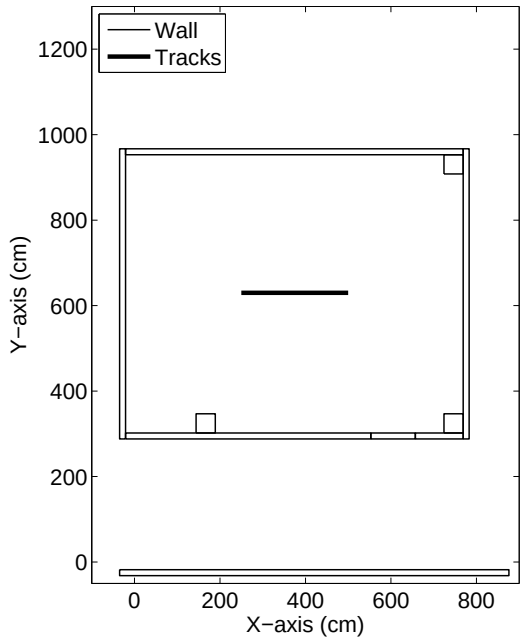


Figure 2.12: Test tracks in the facility.

Recording #	Total Measurements	Speed (cm/s)	Raw Error (cm)
1	1521	11	20
2	467	35	22
3	250	65	23
4	164	100	25
5	135	120	23

Table 2.1: Range of motions

2.2.7 Ground Truth

We use a least squares approach to calculate the ground truth data. The tag is initially placed in the start position and measurements are collected for 15 seconds at this position before moving the tag. After the tag reaches the end position, we wait for another 15 seconds before ending the recording. This gives us "flat" regions near the start and end positions along with linear regions in between which indicate

the movement of the tag. Now, a subset of measurements in each region is used to provide a least squares fit to the set of measurements in that region. The measurement closest to the intersection of the lines determines the start, end or change in dynamics (change in direction of the tag). Figure 2.13 shows a partial output of this approach. In this figure, we can observe the "flat" region indicating the start region and two linear regions indicating the motion of the tag with change in direction.

Since we assume that the tag is moving with a constant velocity, we can associate each measurement with a ground truth location. The velocity is calculated by dividing the total ground truth distance covered by the total time taken when the tag is in motion. Now, multiplying the velocity by the time at which the measurement was received gives us the corresponding ground truth position of the tag at that time instant. This can be written as

$$\breve{x}_t = \frac{D_x}{T} \times t \tag{2.6}$$

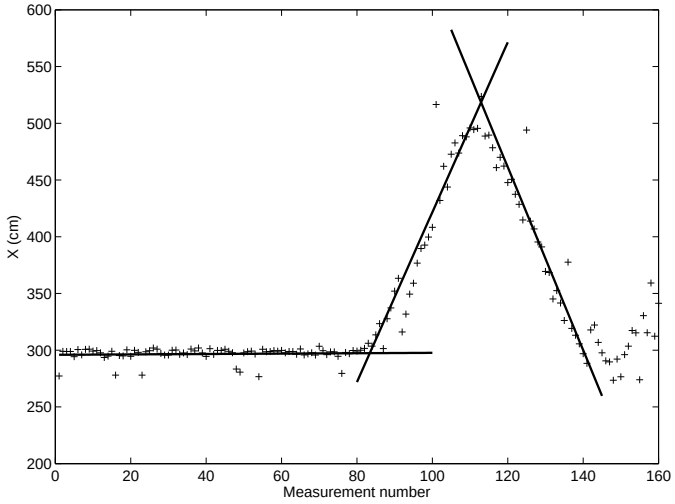
$$\breve{y}_t = \frac{D_y}{T} \times t \tag{2.7}$$

where $\breve{x}_t, \breve{y}_t$ are the ground truth data at time t, D_x, D_y are the total ground truth distances along x and y axes respectively, and T is the total time taken to complete a recording .

2.2.8 Error metric

In order to evaluate the performance of the filtered output we calculate the average Euclidean distance between the filtered data and the corresponding ground truth data over the total number of measurements. This distance is known as the

Figure 2.13: Least squares approach to generate ground truth data (partial output).



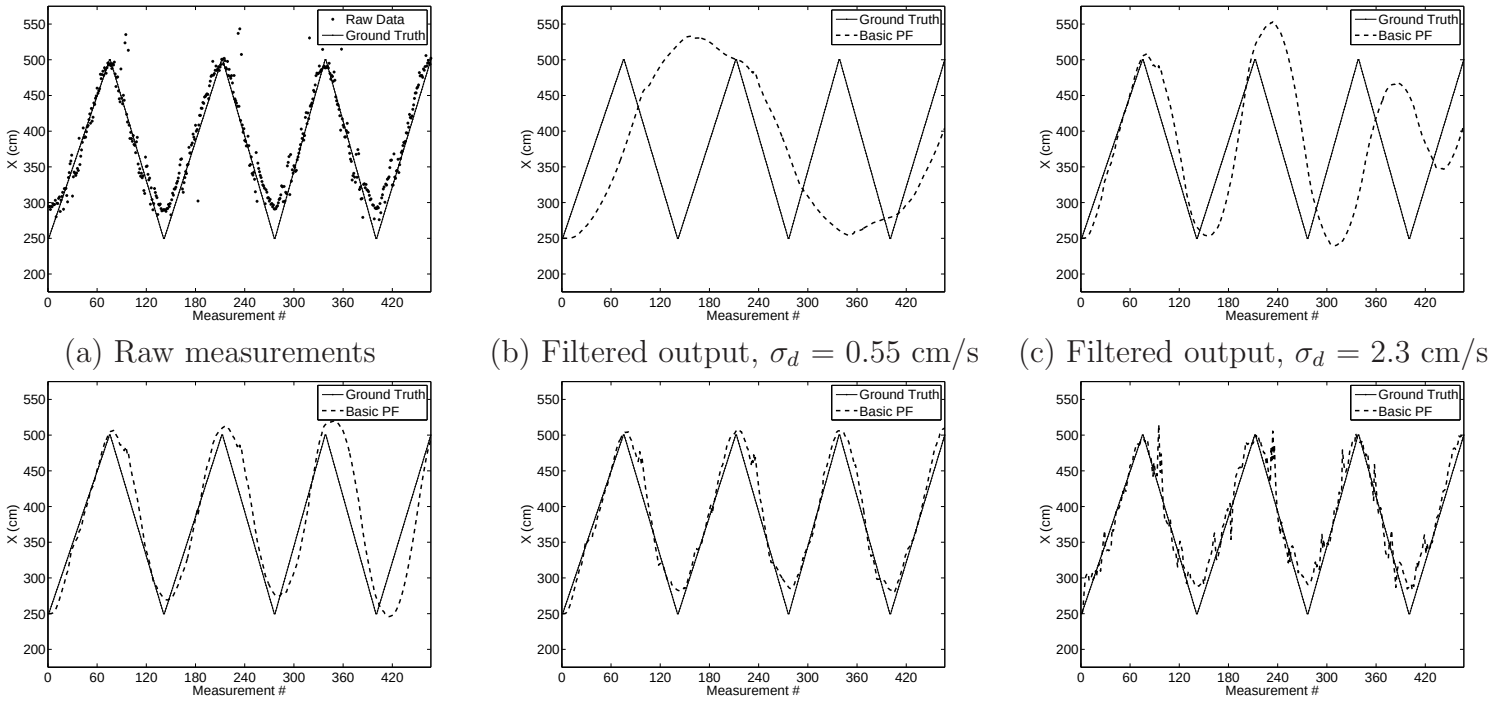
position error (P.E.) and can be defined as

$$P.E. = \frac{1}{N}\sum_{i=1}^{N}\sqrt{(x_i - \breve{x}_i)^2 + (y_i - \breve{y}_i)^2} \tag{2.8}$$

where, x_i, y_i are the filtered data, $\breve{x}_i, \breve{y}_i$ are the corresponding ground truth data for measurement i and N is the total number of measurements.

2.3 Experimental Results

The output for the particle filter is to some degree controlled by the value chosen for σ_d, the dynamic noise in the motion model. This value represents the amount of expected change in velocity at each time step. The lower this value, the more the filter weights the output towards the system equations, in essence providing more smoothing. The higher this value, the more the filter weights the output towards the measurements, allowing a quicker reaction to actual dynamics at the cost of less smoothing. Figure 2.14 demonstrates this effect. Part (a) shows the raw measure-



(a) Raw measurements (b) Filtered output, $\sigma_d = 0.55$ cm/s (c) Filtered output, $\sigma_d = 2.3$ cm/s

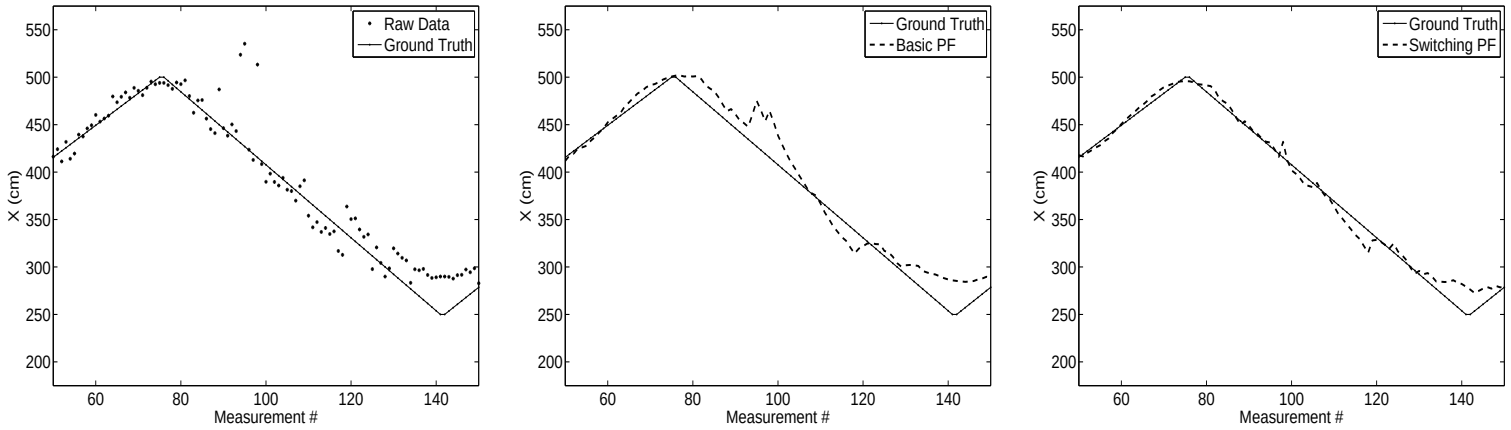
(d) Filtered output, $\sigma_d = 2.8$ cm/s (e) Filtered output, $\sigma_d = 3.7$ cm/s (f) Filtered output, $\sigma_d = 55.0$ cm/s

Figure 2.14: Illustration of the effect of σ_d on filter output.

ments for a recording; parts (b)-(f) show the basic particle filter output for increasing values of σ_d. In part (b), the actual dynamics in the recording (the step changes in velocity) are considerably larger than the value chosen for σ_d, so that the filter output is not able to reliably track the motion. In effect, the filter is smooothing too much. In parts (c)-(d), the larger values for σ_d cause the filter to more reliably track the actual motion, but there is a noticeable lag, particularly at the points where the actual dynamics change. In part (e), σ_d most closely matches the actual dynamics so that the best filter output is obtained. Part (f) shows the output for an even larger value of σ_d, where the filter is giving too much weight to individual measurements.

Figure 2.16 shows the error curves comparing the raw measurements, basic particle filter, and set noise particle filter, for one recording. The error is shown over a range of $\sigma_d = 0.01$ to 60 cm/s. It is important to evaluate performance across a range of σ_d because in practice it is impossible to know the actual dynamics of the motion. The error curves are the average errors of the recordings over 100 trials (repeated runs of each filter at each value of σ_d); this is necessary because the particle filter is a Monte Carlo approach and a single trial of limited length does not necessarily provide a typical representative output. From the figure, it can be seen that the set noise model particle filter performs better than basic particle filter over the entire range of dynamic noise. The minimum error is 15 cm, and occurs at approximately $\sigma_d = 6.0$ cm/s, where σ_d best matches the actual dynamics of the motion for this recording.

Figure 2.15 shows the basic particle filter and set noise particle filter output for this recording at $\sigma_d = 6.0$ cm/s. For clarity, only a subset of the data is shown, and only the X-coordinates are shown (the motion is along a straight line of constant Y). Because the best value of σ_d was chosen for this figure, both filters provide a fairly good output that is better than the raw measurements. However, it can also be seen



(a) Raw measurements (b) X - axis for basic particle filter (c) X - axis for set noise particle filter

Figure 2.15: Recording 2 at $\sigma_d = 6.0$ cm/s (partial output)

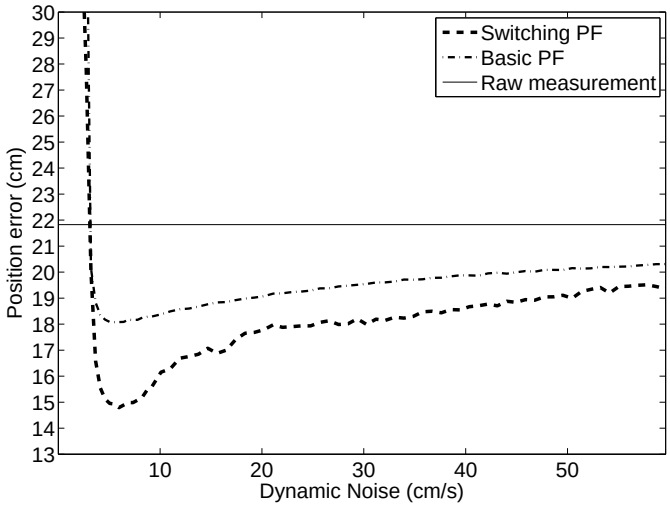


Figure 2.16: Error curve for one recording.

that the set noise particle filter output is more accurate, particularly in the range of measurements from 80 to 120.

Figure 2.17 shows the average error curves for all five recordings. From this figure, it can be observed that the set noise particle filter performs better than the basic particle filter across the entire range of dynamic noise. There is no global minimum like there was in Figure 2.16 because the actual dynamics of all the recordings vary (see Table 2.1). The average accuracy of the raw measurements is approximately 23 cm. The range of $\sigma_d = 30$ to 50 cm/s shows that the set noise particle filter improved the accuracy of the raw measurements by approximately 4 cm on average, about double that of basic particle filter. Thus, our set noise particle filter shows an approximately 15% improvement over a basic particle filter.

2.4 Conclusions

We have identified a new noise source due to the switching of fixed point sets for trilateration. While this noise is theoretically present in all trilateration-based

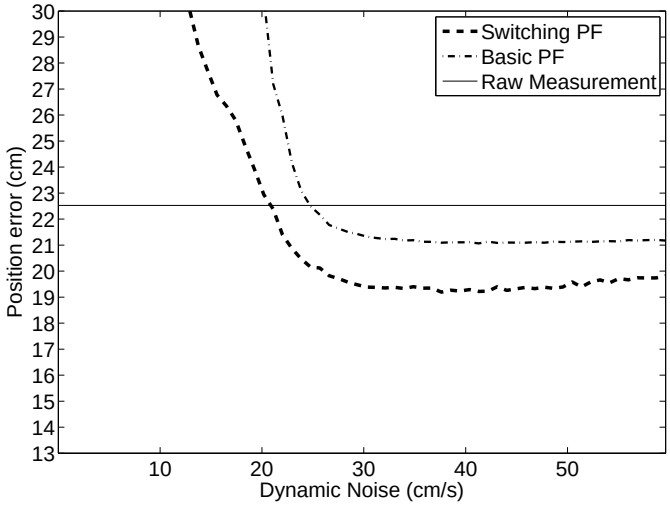


Figure 2.17: Average error curve for all recordings.

systems, it is not readily apparent in large-scale systems like a GNSS, but it can cause noticeable jump-like behavior in indoor UWB position tracking. We have developed a mathematical model and particle filter that accounts for this noise. We tested our methods on a real UWB indoor position tracking system. Our set noise particle filter showed an approximately 15% improvement in accuracy over the raw measurements. We use a particle filtering approach since it can handle non-zero mean noise models, unlike Kalman filtering approaches. In the experiments, we have considered Gaussian noise models for the sensor sets. However, this work can be easily extended to use non-Gaussian noise models.

Our experiments have been conducted in a real world setting, where we have achieved modest improvement in the tracking accuracy over a range of dynamics. However, we have observed that noise due to NLOS, multipath and timing errors tend to be more significant than noise caused by sensor set switching. Hence, research studying the impact of these noise sources in isolation have been able to achieve sub-decimeter [96, 99] and sometimes even sub-centimeter accuracies [100, 160], while we have been able to observe only a modest improvement.

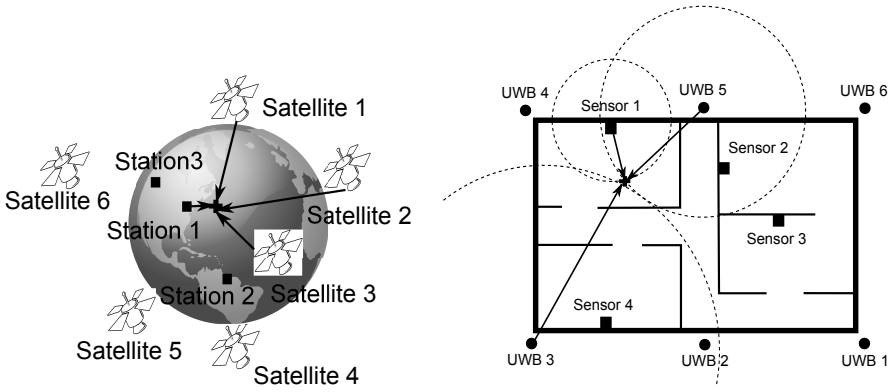
Chapter 3

Augmentation

3.1 Motivation

The work in this chapter is inspired by methods used to improve the accuracy of a GNSS through augmentation. Augmentation methods use additional instrumentation such as signals from a nearby differential broadcast station, maps or onboard inertial systems [52]. Figure 3.1 (a) shows an example of a differential GNSS (DGNSS) where the satellites are used to provide a rough estimate of the location and the broadcast stations provide correctional signals to improve the precision. Since the correctional signals are broadcast from a location which is much closer than the satellites, the measured distances are more precise. Moreover, these signals are not affected by noise due to changes in permittivity in the atmosphere and relativistic effects. Therefore these correctional signals are used to augment the raw GNSS measurements.

Since the principle of operation used in UWB indoor position tracking is similar to that used in a GNSS, we propose to explore if the same approach of using a differential signal can be used to improve accuracy. Figure 3.1 (b) shows a diagram



(a) Differential GNSS. (b) Differential LPS.

Figure 3.1: Example of augmentation on a global and local level.

of how a UWB positioning system could be augmented by using additional sensors in each room. We assume that the UWB sensors placed around the facility provide a rough estimate of the position of the target. Since the UWB sensors are NLOS, the position estimates are less precise. A differential sensor placed in each local area (here, room level) behaves similar to the broadcast tower in a DGNSS. Since these differential sensors are placed LOS, they may be able to provide measurements with higher precision. These higher precision estimates can then be used to augment the estimates from the UWB sensors.

In our experiments we test two types of augmentation sensors, a LADAR and a network of cameras. The LADAR provides a 1D waist-high distance scan of the room. The camera network provides a floor-level image of occupied space (known as an occupancy map [66]). Fusing each of these sensors with a UWB system introduces an interesting challenge in data fusion. The UWB system requires tracked targets to carry a tag, thus actively aiding in their tracking. Each tag transmits a unique code, identifying the target and simplifying data association over time. Both the LADAR and camera network do not have this advantage, and instead rely upon segmentation and temporal data association to determine and maintain track identities. When

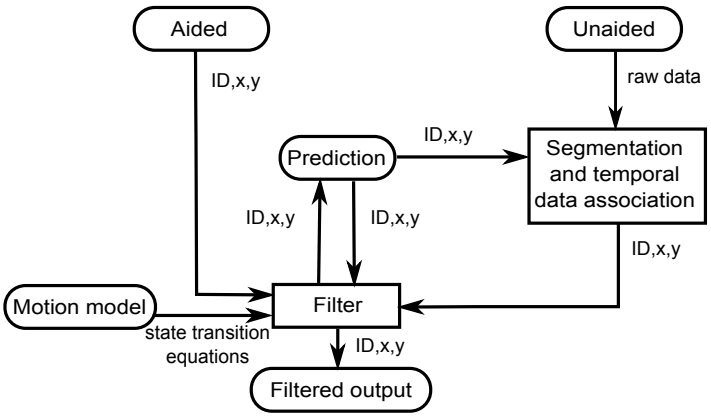


Figure 3.2: Aided and unaided fusion framework.

only one person is in a room, these problems are relatively trivial, but as the number of people increases, the segmentation and data association problems become more difficult for the LADAR and camera network. We therefore introduce a new data fusion framework that combines data from aided and unaided sensors.

The main goal of the work in this chapter is to explore the potential for a differential augmentation to improve UWB indoor position tracking. However, due to the interesting nature of the data fusion problem we first describe a new data fusion framework. We then discuss a switching observation model Kalman filter that is used to combine asynchronous measurements received from multiple sensors. To evaluate our methods we use error metrics developed by the tracking research community that specifically address the two problems of improving precision and data association. Finally, we demonstrate the operation of the framework on data from a real system and discuss the results obtained from the fusion.

3.2 Methods

3.2.1 Framework

We propose a general algorithm to combine asynchronous measurements from aided and unaided sensors in a filtering framework. For an aided sensor the data association problem we assume that the targets broadcast their identity in addition to sensor readings. For an unaided sensor, we assume that the sensor data must be segmented and associated with tracking identities. Figure 3.2 shows a diagram of our framework. There are three sources of input: aided data, unaided data, and a motion model. For an aided sensor such as a UWB sensor, we assume that the ID is provided uniquely by the sensor for every tracked target in addition to its x,y position. For an unaided sensor such as a LADAR or a camera network, we assume that the raw data needs to be segmented to identify the targets. Once the positions of the targets have been identified they have to be associated. The association can be done temporally by using the previous positions or by using identity information from another source. The aided and unaided measurements are combined using a filter such as a Kalman or particle filter. The filter requires some prior knowledge of the target's motion model. The motion model is used to provide an estimate of where the target is expected to move based upon state transition equations. In our experiments we use a discrete Kalman filter. A detailed explanation of the working of the discrete Kalman filter is provided in section 1.6.2.

Figure 3.3 shows the flow of data in the proposed fusion framework. The aided sensor provides the ID and x, y positions of the target which is directly fed to the filter. The filter uses the received measurement to update the estimate provided by the motion model. Based on the motion model the filter makes a prediction which is fed back to the filter if no measurement from the unaided sensor is available.

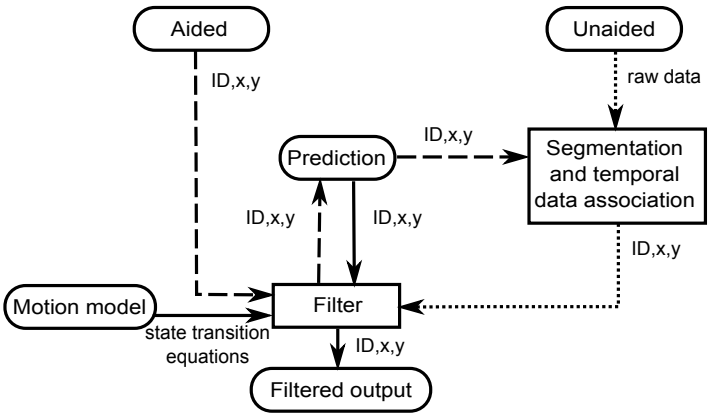


Figure 3.3: Working of an aided and unaided fusion system.

If a measurement from the unaided sensor is available then the prediction is used to segment the raw data and assist in data association. We used a nearest neighbor approach to associate the predicted data with the segmented data, but other methods could also be used. The associated data is fed to the filter which is then used to update the filter prediction. The updated state is then used as the filtered output.

3.2.2 Facility

Our facility is located in the basement of Riggs Hall at Clemson University. The test area covers approximately 8 m $\times$ 8 m, covering the majority of the sensor network laboratory and part of an adjacent hallway. Figure 3.4 shows the locations of furniture, walls, and the UWB, LADAR and camera sensors in the test area. A detailed description of the facility is provided in section 2.2.2. Since the LADAR and camera sensor can operate only under LOS conditions we conduct experiments only in this one room where all sensors are present.

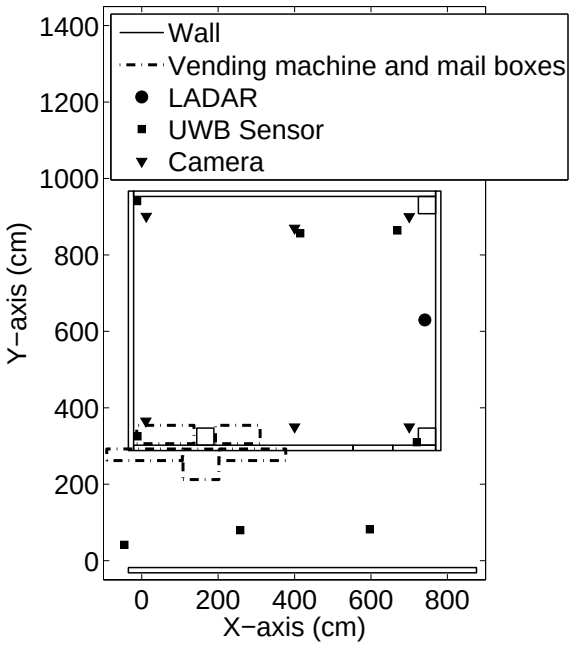


Figure 3.4: Positions of UWB, LADAR and camera sensors.

3.2.3 Sensors

The UWB positioning system is described in section 2.2.3. The LADAR and camera systems are connected to the same computer, called 'Fusion-Net', while the UWB positioning system is connected to a standalone computer 'Ubi-Net' and forwards data to the 'Fusion-Net' computer. Three separate threads are used to run these sensor systems. The time stamped data arriving from the sensors are stored on the computer for offline analysis. The operation of the sensors are shown using a block diagram in Figure 3.5. Sections 3.2.3.1 and 3.2.3.2 briefly describe the LADAR and camera systems and their operation.

3.2.3.1 LADAR system

In our experiments we used a commercially available SICK LMS 291-S05 LADAR manufactured by SICK AG. This sensor is shown in Figure 3.6. The LADAR

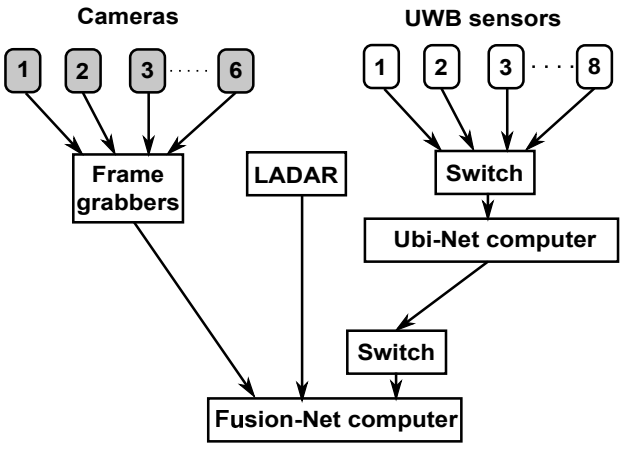
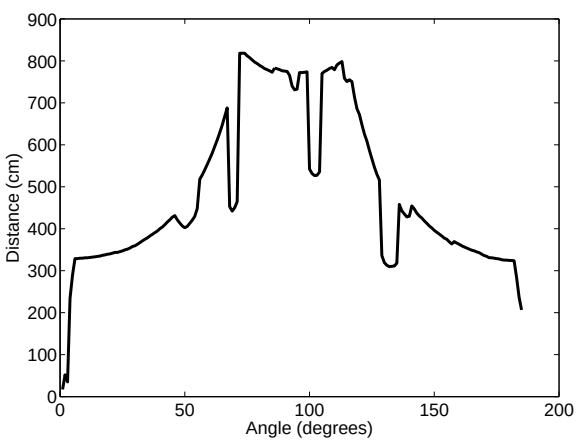


Figure 3.5: Block diagram showing the connection of the sensor systems.

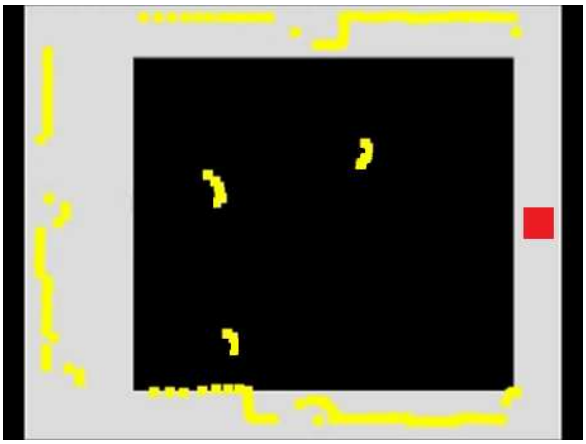
emits a beam of infrared signal across the tracking space. It then calculates the time taken by the emitted laser beam to return after being reflected from an object [131]. Figure 3.7 shows the LADAR in operation and its raw measurements. In Figure 3.7 the x-axis corresponds to the angle while the y-axis corresponds to the distance from the LADAR. In this case there are 3 objects present in the tracking area that can be seen by the shorter distances in the LADAR measurements. The red box corresponds to the LADAR which scans 0 to 180 degrees from top to bottom in a counter-clockwise fashion. The yellow squares correspond to the distance measurements at each angular resolution. In the figure, the raw distance and angular measurements have been converted into x,y measurements using simple trigonometric equations.

The LADAR sensor has a typical range of 30m with a range resolution of 10mm/$\pm$35mm and angular resolutions of 0.25°, 0.5° and 1°. In our experiment, we set the scan range to 50 m with an angular resolution of 1°. The sensor system operates at approximately 2.5 Hz. The position of the LADAR is shown in figure 3.4. Its position has been carefully measured to within 1 cm using knowledge of the position of the center of the mirror wheel obtained from its schematic diagram (see Figure 3.8) and the global coordinate system. The LADAR has been placed at a

Figure 3.6: SICK LADAR - LMS291.



(a) In d-θ space



(b) In x-y space

Figure 3.7: Raw measurements from a LADAR.

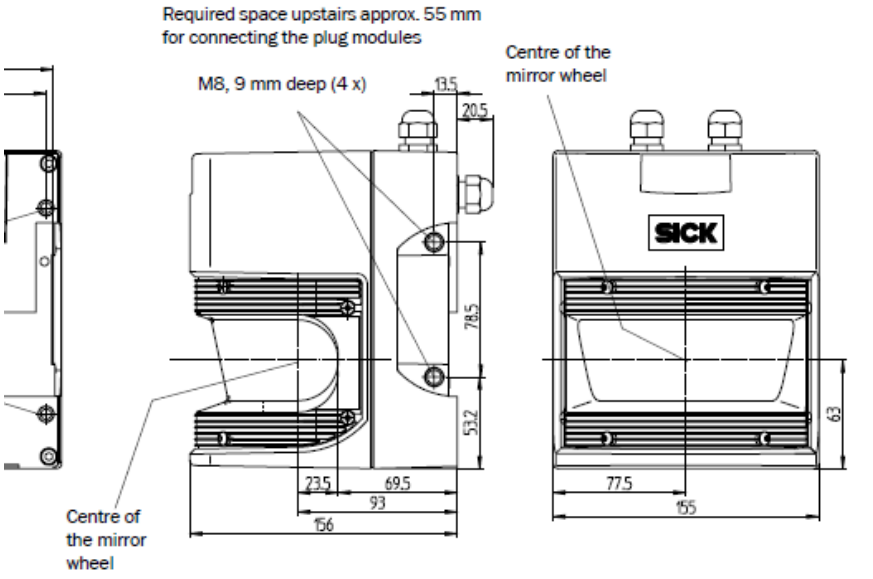


Figure 3.8: Schematic diagram of LMS291.

height of approximately 95 cm to correspond to the waist height of an average person. A picture of the sensor in the room is shown in figure 3.9.

3.2.3.2 Camera system

We use a network of six video cameras and fuse their intensity data to create an occupancy map that is able to track people in a x-y space [66]. The cameras are placed around the room as shown in figure 3.4, pointed towards the center of the room. The cameras operate at 30 Hz generating an occupancy map at a rate of roughly 20 Hz. The occupancy map is a 480 $\times$ 640 image where each pixel indicates the space in the plane of the floor is either empty or occupied. Figure 3.11 shows an example of an occupancy map. In this example, a single person is walking around in the room and can be seen as a blob. A blob detection algorithm is used to track the motion of people in the occupancy map. The output from the algorithm is the x,y position of the centroids of the blobs. These positions are then converted from the

Figure 3.9: Picture of LADAR in the room.

occupancy map space to the UWB coordinate space.

3.2.4 Filtering

Section 1.6.2 describes the basic Kalman filter. For this work, we use a variation of the Kalman filter to fuse observations received from two or more sensors. The state transition equations remain the same as described in section 1.6.1, but for observations our filter uses a switching observation model [39, 154]. Since the sensors operate asynchronously, there are three possibilities of the types of data that are available for a filter iteration, as shown in figure 3.12.

For the fusion between UWB and LADAR the three cases are enumerated in equation 3.1. In the first case, a measurement is received from only the aided sensor, here UWB. In this case the observations, observation matrix and corresponding noise covariance given by equations 1.9, 1.10 and 1.14 are replaced by those specific to the UWB sensor. In the second case, when data is received from only the LADAR then these matrices are replaced accordingly. In the third case when measurements are

Figure 3.10: Figure showing cameras placed in the room.

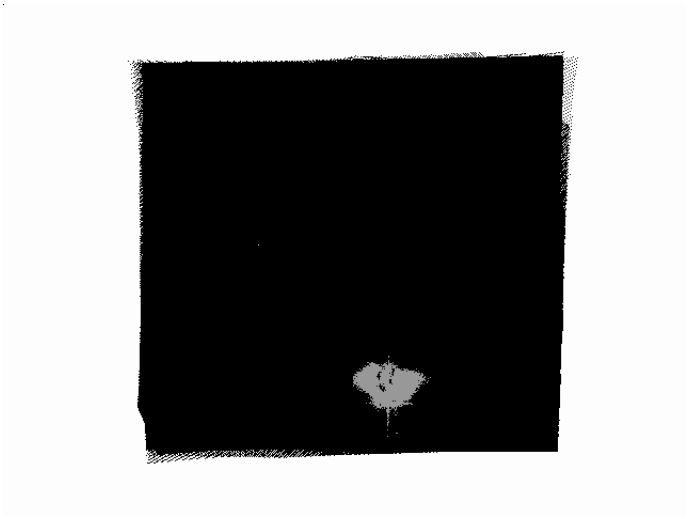


Figure 3.11: Example of an occupancy map.

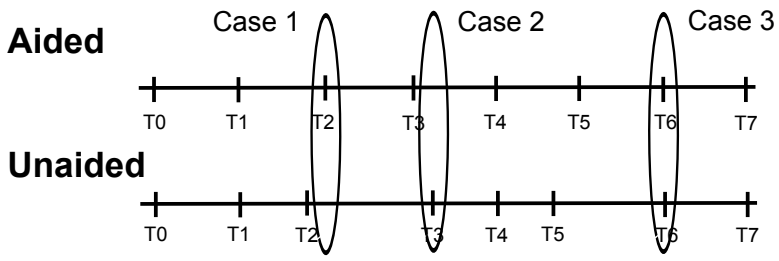


Figure 3.12: Illustration of timing behavior in asynchronous fusion framework.

received from both UWB and LADAR simultaneously then the observation and noise covariance matrices are modified accordingly.

$$\mathbf{Z}, \mathbf{O}, \mathbf{R} = \begin{cases} \mathbf{Z}_u, \mathbf{O}_u, \mathbf{R}_u & : \text{when only UWB is present} \\ \mathbf{Z}_l, \mathbf{O}_l, \mathbf{R}_l & : \text{when only LADAR is present} \\ \mathbf{Z}_{ul}, \mathbf{O}_{ul}, \mathbf{R}_{ul} & : \text{when both are present} \end{cases} \tag{3.1}$$

Similarly, for the fusion between UWB and the camera network, the $\mathbf{Z}$, $\mathbf{O}$ and $\mathbf{R}$ matrices are replaced based on the type of measurement received at the given time.

$$\mathbf{Z}, \mathbf{O}, \mathbf{R} = \begin{cases} \mathbf{Z}_u, \mathbf{O}_u, \mathbf{R}_u & : \text{when only UWB is present} \\ \mathbf{Z}_c, \mathbf{O}_c, \mathbf{R}_c & : \text{when only camera is present} \\ \mathbf{Z}_{uc}, \mathbf{O}_{uc}, \mathbf{R}_{uc} & : \text{when both are present} \end{cases} \tag{3.2}$$

The $\mathbf{Z}$, $\mathbf{O}$ and $\mathbf{R}$ matrices for the UWB sensor can be written as

$$\mathbf{Z}_u = \begin{bmatrix} \tilde{x}_{ut} \\ \tilde{y}_{ut} \end{bmatrix} \tag{3.3}$$

$$\mathbf{O}_u = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{3.4}$$

$$\mathbf{R}_u = \begin{bmatrix} \sigma^2_{n_{ux}} & \sigma_{n_{ux},n_{uy}} \\ \sigma_{n_{ux},n_{uy}} & \sigma^2_{n_{uy}} \end{bmatrix} \tag{3.5}$$

The $\mathbf{Z}$, $\mathbf{O}$ and $\mathbf{R}$ matrices for the LADAR sensor can be written as

$$\mathbf{Z}_l = \begin{bmatrix} \tilde{x}_{lt} \\ \tilde{y}_{lt} \end{bmatrix} \tag{3.6}$$

$$\mathbf{O}_l = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{3.7}$$

$$\mathbf{R}_l = \begin{bmatrix} \sigma^2_{n_{lx}} & \sigma_{n_{lx},n_{ly}} \\ \sigma_{n_{lx},n_{ly}} & \sigma^2_{n_{ly}} \end{bmatrix} \tag{3.8}$$

The $\mathbf{Z}$, $\mathbf{O}$ and $\mathbf{R}$ matrices for the camera network can be written as

$$\mathbf{Z}_c = \begin{bmatrix} \tilde{x}_{ct} \\ \tilde{y}_{ct} \end{bmatrix} \tag{3.9}$$

$$\mathbf{O}_c = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{3.10}$$

$$\mathbf{R}_c = \begin{bmatrix} \sigma^2_{n_{cx}} & \sigma_{n_{cx},n_{cy}} \\ \sigma_{n_{cx},n_{cy}} & \sigma^2_{n_{cy}} \end{bmatrix} \tag{3.11}$$

When observations from both UWB and LADAR sensors are available at the same time, the matrix $\mathbf{Z}$ is adjusted to include both observations and is written as

$$\mathbf{Z}_{ul} = \begin{bmatrix} \tilde{x}_{ut} \\ \tilde{y}_{ut} \\ \tilde{x}_{lt} \\ \tilde{y}_{lt} \end{bmatrix} \tag{3.12}$$

where $\tilde{x}_{ut}, \tilde{y}_{ut}$ are the observations obtained from UWB and $\tilde{x}_{lt}, \tilde{y}_{lt}$ are the observations obtained from the LADAR.

The observation equations g are

$$\boldsymbol{g} = \begin{bmatrix} \tilde{x}_{ut} = x_{ut} + \mathcal{N}(0, \sigma_{n_{ux}}) \\ \tilde{y}_{ut} = y_{ut} + \mathcal{N}(0, \sigma_{n_{uy}}) \\ \tilde{x}_{lt} = x_{lt} + \mathcal{N}(0, \sigma_{n_{lx}}) \\ \tilde{y}_{lt} = y_{lt} + \mathcal{N}(0, \sigma_{n_{ly}}) \end{bmatrix} \tag{3.13}$$

where $\mathcal{N}(0, \sigma_{n_{ux}})$ and $\mathcal{N}(0, \sigma_{n_{uy}})$ is the 2D zero-mean Gaussian noise for UWB, and $\mathcal{N}(0, \sigma_{n_{lx}})$ and $\mathcal{N}(0, \sigma_{n_{ly}})$ describes the 2D zero-mean Gaussian noise of the LADAR.

The observation matrix can be written as

$$\mathbf{O}_{ul} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{3.14}$$

The measurement noise covariance matrix $\mathbf{R}$ is rewritten as

$$\mathbf{R}_{ul} = \begin{bmatrix} \sigma^2_{n_{ux}} & \sigma_{n_{ux},n_{uy}} & 0 & 0 \\ \sigma_{n_{ux},n_{uy}} & \sigma^2_{n_{uy}} & 0 & 0 \\ 0 & 0 & \sigma^2_{n_{lx}} & \sigma_{n_{lx},n_{ly}} \\ 0 & 0 & \sigma_{n_{lx},n_{ly}} & \sigma^2_{n_{ly}} \end{bmatrix} \tag{3.15}$$

Similarly when observations from both UWB and camera sensors are available,

the matrix $\mathbf{Z}$ is written as

$$\mathbf{Z}_{uc} = \begin{bmatrix} \tilde{x}_{ut} \\ \tilde{y}_{ut} \\ \tilde{x}_{ct} \\ \tilde{y}_{ct} \end{bmatrix} \tag{3.16}$$

where $\tilde{x}_{ut}, \tilde{y}_{ut}$ are the observations obtained from UWB and $\tilde{x}_{ct}, \tilde{y}_{ct}$ are the observations obtained from the camera.

The observation equations g are

$$\boldsymbol{g} = \begin{bmatrix} \tilde{x}_{ut} = x_{ut} + \mathcal{N}(0, \sigma_{n_{ux}}) \\ \tilde{y}_{ut} = y_{ut} + \mathcal{N}(0, \sigma_{n_{uy}}) \\ \tilde{x}_{ct} = x_{ct} + \mathcal{N}(0, \sigma_{n_{cx}}) \\ \tilde{y}_{ct} = y_{ct} + \mathcal{N}(0, \sigma_{n_{cy}}) \end{bmatrix} \tag{3.17}$$

where $\mathcal{N}(0, \sigma_{n_{ux}})$ and $\mathcal{N}(0, \sigma_{n_{uy}})$ is the 2D zero-mean Gaussian noise for UWB, and $\mathcal{N}(0, \sigma_{n_{cx}})$ and $\mathcal{N}(0, \sigma_{n_{cy}})$ describes the 2D zero-mean Gaussian noise of the camera.

The observation matrix can be written as

$$\mathbf{O}_{uc} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{3.18}$$

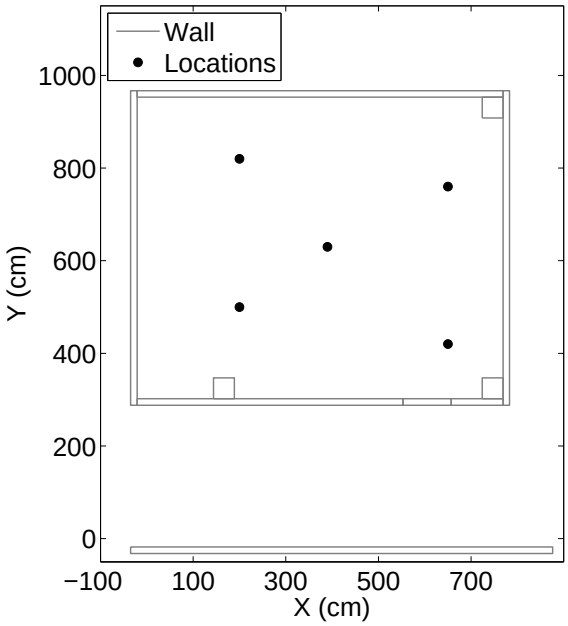


Figure 3.13: Calibration locations for unaided sensors.

The measurement noise covariance matrix $\mathbf{R}$ is rewritten as

$$\mathbf{R}_{uc} = \begin{bmatrix} \sigma^2_{n_{ux}} & \sigma_{n_{ux},n_{uy}} & 0 & 0 \\ \sigma_{n_{ux},n_{uy}} & \sigma^2_{n_{uy}} & 0 & 0 \\ 0 & 0 & \sigma^2_{n_{cx}} & \sigma_{n_{cx},n_{cy}} \\ 0 & 0 & \sigma_{n_{cx},n_{cy}} & \sigma^2_{n_{cy}} \end{bmatrix} \tag{3.19}$$

3.2.4.1 Parameters

The noise model for the UWB sensor ($\sigma_{n_{ux}}$ and $\sigma_{n_{uy}}$) was set to 40 cm based on previous experiments conducted using the same system in the same facility [139]. In order to calculate the standard deviation in the measurements for the LADAR and camera sensors a calibration was conducted. A circular, metallic can having a diameter of 40 cm which is roughly close to the diameter of an average person was placed at 5 different locations in the room (Figure 3.13). At each of these locations 100

measurements for the LADAR and 800 measurements for the camera were collected. The error in the measurements from the LADAR and camera network with respect to the ground truth were calculated in addition to the standard deviation in the measurements. For the LADAR, the values of standard deviations were calculated as $\sigma_{n_x} = 0.28$ cm and $\sigma_{n_y} = 0.16$ cm. For the camera network, values of $\sigma_{n_x} = 0.27$ cm and $\sigma_{n_y} = 0.15$ cm were calculated. In both these cases, the observation equations are the same as described by equation 1.10.

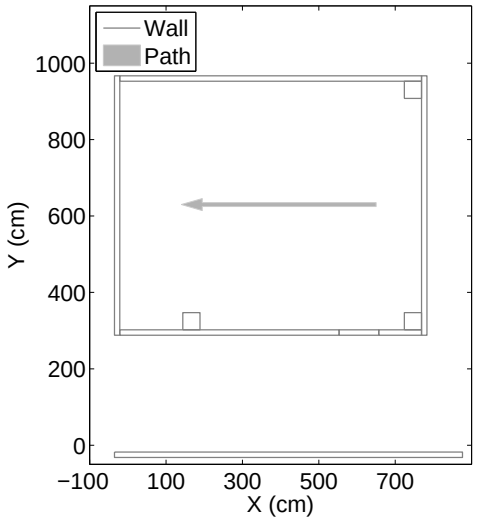


Figure 3.14: Path at which recordings were conducted to determine σ_d.

As mentioned in section 2.3, the choice of σ_d in the filter has some control over the output, defining the expected nominal value of acceleration. To determine the best value of dynamic noise for the fusion we collected 7 recordings in the center of the room with a person walking along a straight line as shown in Figure 3.14. The person walked at different speeds ranging from roughly 14 to 75 cm/s. We plotted the combined average error of UWB with LADAR, and UWB with camera sensors for these recordings at different values of σ_d ranging from 0.01 to 1.5 cm/s (Figure 3.15). From the figure we can observe that the knee of the curves is near 0.2 cm/s.

We therefore choose the value of $\sigma_d = 0.2$ cm/s in our experiments.

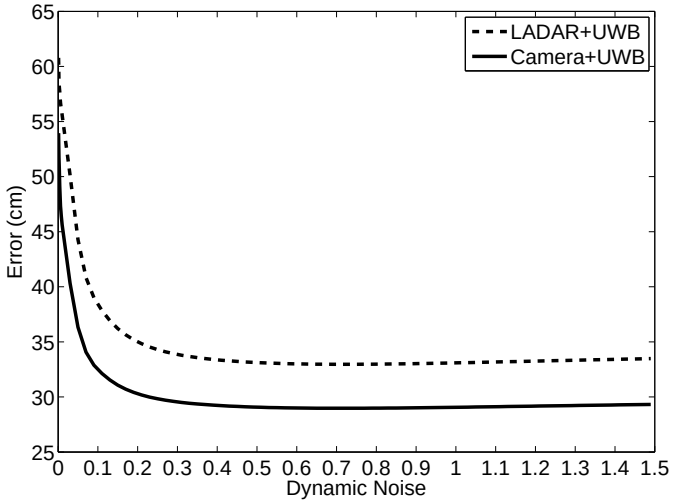


Figure 3.15: Average error curves for choosing σ_d.

3.2.5 Data Collection

We used laser levels and tape measures to ensure that the ground truth locations are accurate to within 1 cm relative to the calibrated Ubisense coordinate system. The location of the LADAR was also carefully surveyed and aligned to the Ubisense coordinate system. The cameras have also been calibrated to the same coordinate system. Eight different paths were laid on the ground as shown in Figure 3.16. The arrows indicate the direction of motion taken by a person along each path. We varied the number of people present in the room from 1 to 8. For the cases where the number of people was varied from 1 to 3 we collected 10 recordings for a total of 30 recordings. Where the number of people was varied from 4 to 8 we collected 5 recordings for a total of 25 recordings. Table 3.1 shows the start and end positions of each path.

We placed the UWB tag on a helmet which was worn by the people being tracked. This was done so as to provide a rough approximate of the center of the

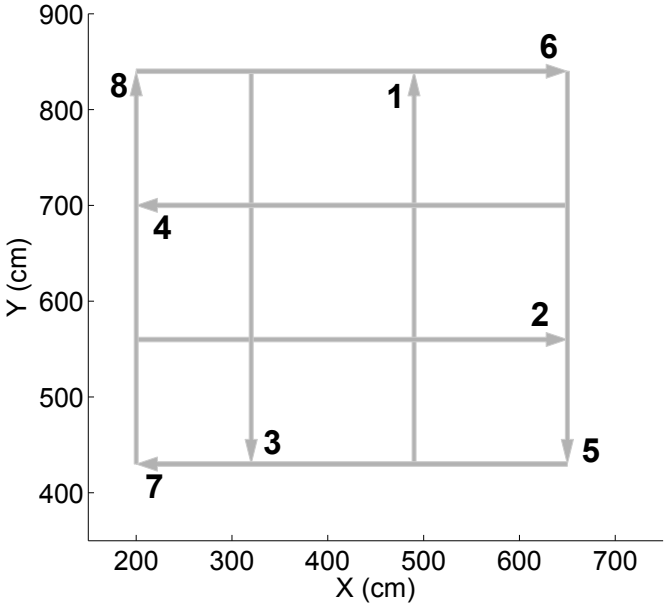


Figure 3.16: Test paths in the facility.

Path #	Start (cm)	End (cm)
1	(490,430)	(490,840)
2	(200,560)	(650,560)
3	(320,840)	(320,430)
4	(650,700)	(200,700)
5	(650,840)	(650,430)
6	(200,840)	(650,840)
7	(650,430)	(200,430)
8	(200,430)	(200,840)

Table 3.1: Start and end positions of the paths.

human body. Figure 3.17 shows a picture of tracking 8 people during a recording. We assumed that the subjects are moving at a relatively constant speed and as accurately as possible along the path. We also constrained the speed at which people move to not more than 30 cm/s. The reason for this constraint is that the LADAR operates

Figure 3.17: A snapshot of data collection using aided and unaided sensors.

at only 2.5 Hz. This rate is not fast enough to adequately sample the motion of a normal human pace (which is $\approx$ 100 cm/s). Due to the complexity of evaluating a multi-object tracking system for multiple sensors with respect to a known ground truth we consider only linear tracks.

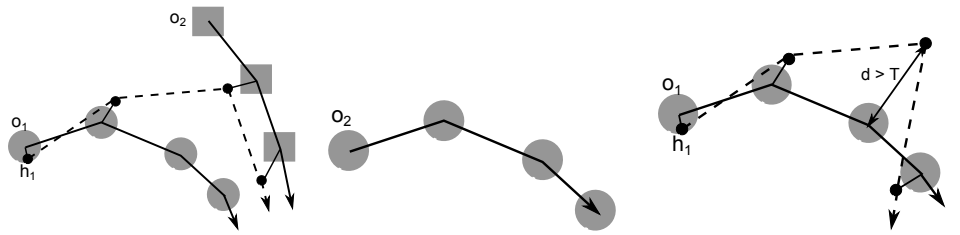
3.2.6 Error metrics

For single person tracking systems where the target's path is exactly known the performance of an algorithm can be evaluated based on its precision. Here *precision* is defined as the average distance between the ground truth and actual positions of the target over all measurements. This metric however does not evaluate potential data association errors. In a multi-target tracking system the targets may interact with each other causing occlusions that lead to missed data, data association errors, and false positives.

The Performance Evaluation of Tracking and Surveillance (PETS) workshop [118] was convened to help develop standardized metrics for evaluating the performance of object detection and tracking algorithms. Ellis [36] categorized metrics that

have been widely used by researchers to quantify the performance of a system into two categories – statistical methods and numeric scores. The first category consists of techniques that use standard statistical methods to compare two population values while the second category consists of techniques that calculate numeric values such as the accuracy of detection, average positional errors, specificity, sensitivity, positive predictive value, false positive and negative and negative predictive value. Black et al. [12] proposed metrics based on the calculation of a contingency table which provides a count of the true positives, true negatives, false positives and false negatives. They measured the tracking accuracy by calculating the average Euclidean distance between the ground truth and the measurements. ETISEO (Evaluation du Traitment et de L'Interpretation de Séquences Video), a project sponsored by the French government also proposed a few metrics to evaluate the performance of video surveillance systems [113]. A drawback of all these algorithms is that they quantify the performance of an algorithm over multiple metrics. This makes comparison of different algorithms difficult. To solve this issue, the Computers in the Human Interaction Loop (CHIL) [137, 149] and Video Analysis and Content Extraction (VACE) [72, 101] programs were instituted to develop a spatio-temporal metric which can take into account the number of objects detected and tracked, missed objects, and false positives. The CLEAR workshop was the first international workshop that brought the CHIL and VACE efforts together [22, 135, 136]. The Multi-Object Tracking (MOT) metrics defined by Bernardin et al. [8, 9] address the issues mentioned previously and provide an intuitive quantitative analysis of multiple target trackers based on their spatial and temporal accuracy.

In our work we use the MOT metrics to evaluate the performance of multiple target tracking. At each time step t, the tracker is assumed to provide a set of measurements $\{h_1 \dots h_m\}$ for a set of visible targets $\{o_1 \dots o_n\}$. At each of these time



(a) Mismatch error (b) Miss error (c) False positive error

Figure 3.18: Correspondence error in multi-target tracking system.

steps, we determine the best correspondence between the measurements and ground truth. Once the correspondence has been achieved we calculate the position error. Finally, we calculate the number of correspondence errors including mismatches, misses and false positives. Figure 3.18 shows the three correspondence errors considered in this metric. A correspondence is considered to be a *mismatch* if the measurement correspondence has changed as compared to the previous time step. From Figure 3.18 (a), it can be observed that a set of measurements h_1 is initially associated with target o_1. However, after being associated with ground truth o_1 for the first two steps, the measurements are closer to ground truth o_2. Since ground truth o_2 does not have any valid measurement associated with it, the identity is switched causing a mismatch. If there is no measurement corresponding to the ground truth then it is considered to be a *miss* (see Figure 3.18 (b)). A *false positive* is considered to be the case when a measurement is associated with no visible ground truth or when the measurement is greater than a specified threshold from the ground truth. The latter case is shown in Figure 3.18 (c). At the third time instant, the measurement is greater than a threshold T and is therefore not associated with the ground truth. Since the measurement is not associated with any ground truth it is considered to be a false positive. The threshold T puts an upper limit on the precision error for multi-target tracking. In our experiments we choose T=100 cm based upon a desired maximum error for tracking the position of a person moving around a room.

The total average position error between correctly matched ground truth-measurement pairs over all frames, multiple object tracking precision (MOTP) is calculated as

$$MOTP = \frac{\sum_{i,t} d_t^i}{\sum_t c_t} \tag{3.20}$$

where d_t^i is the error between each target-measurement, and c_t is the number of correct correspondences.

The multiple object tracking accuracy (MOTA) is calculated as

$$MOTA = 1 - \frac{\sum_t (m_t + fp_t + mme_t)}{\sum_t g_t} \tag{3.21}$$

where m_t, fp_t, and mme_t are the number of misses, false positives, and mismatches respectively while g_t is the total number of ground truth targets at time t.

Figure 3.19 illustrates these different errors that can be present in a multi-object tracking system. We consider two people labeled 1 and 2 moving along paths that have been labeled A and B respectively. When the people move close to each other their identities can get switched causing a mismatch error. From the figure we can observe that person 1 was associated with path A and person 2 was associated with path B. After they moved together, person 1 was associated with path B and person 2 was associated with path A. This shows that both persons switched their identities causing two mismatch errors. In this illustration there are 4 false positives present. False positives can occur due to limitations of the sensor and they are shown as stars in the figure. Sensor limitations can also lead to missing data as shown by the circled region.

In sections 3.2.7.1 and 3.2.7.2 we describe simulations conducted to demonstrate these metrics for LADAR and camera systems respectively. Theoretically, as

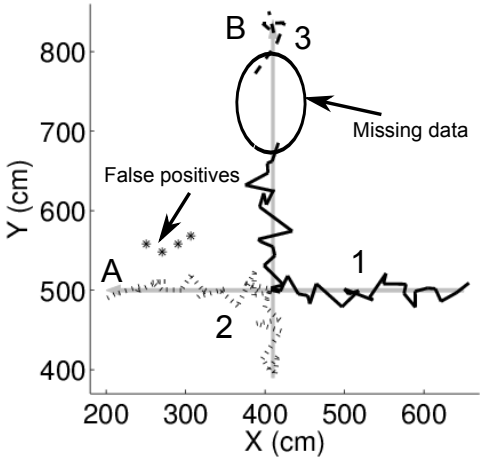


Figure 3.19: Illustration of metrics used in the evaluation of multi-object tracking.

the number of targets in a given tracking area increase there should be a decrease in the performance of the system. This is due to an increase in occlusions/overlaps caused by the interaction of multiple targets, data association issues and other correspondence errors. For sake of simplicity we assume that measurements are received from the sensors synchronously.

3.2.7 Simulations

3.2.7.1 LADAR

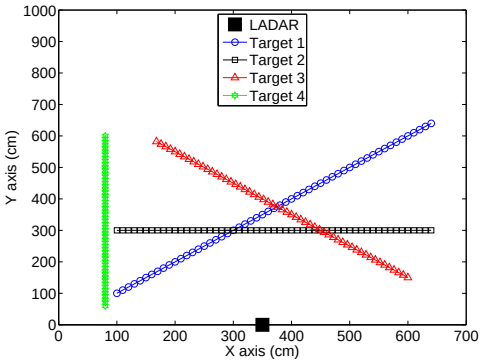
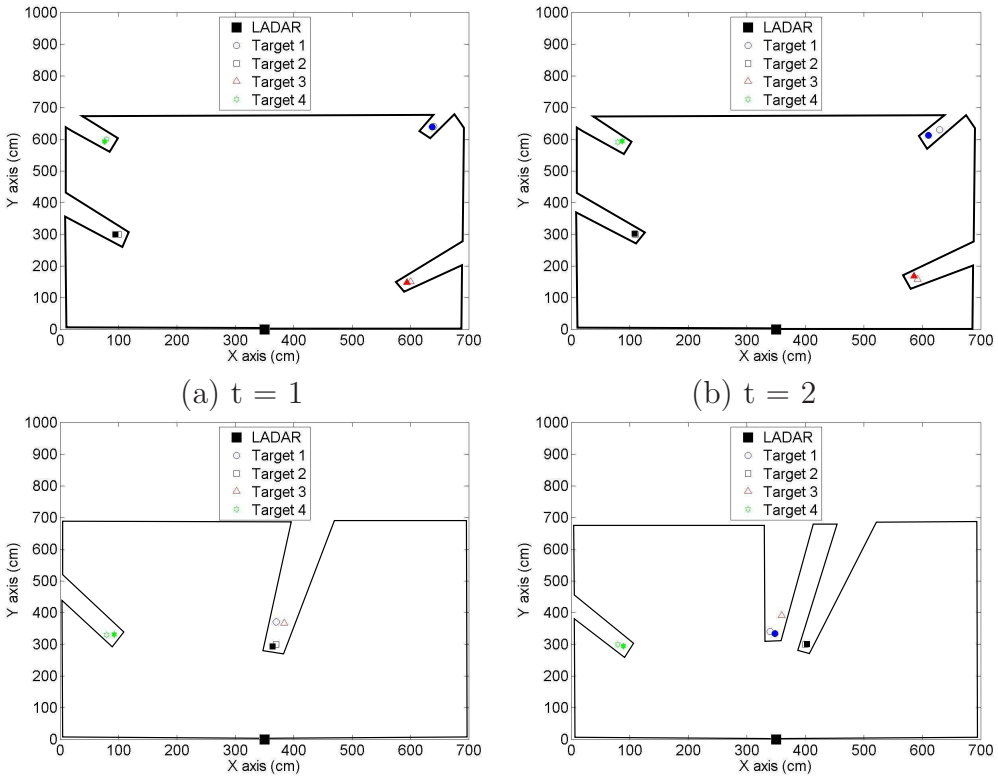


Figure 3.20: Simulated tracks for LADAR based tracking system.

A room 700 x 700 cm^2 in dimension is considered which contains 4 simulated tracks as shown in Figure 3.20. We have added a Gaussian noise with a standard deviation of 10 cm to the ground truth positions to create measurements. The LADAR is assumed to be placed at (X,Y) = (350,0) cm facing the Y-axis and scans counter-clockwise from 0° – 180°. The objects move in a 2D X-Y space as shown in the figure. We assume that the tracking is conducted over 55 time steps. The LADAR provides only the distance and the angle to the target as measurements which is then converted to X-Y coordinates in order to display its position in the figure. The LADAR being an unaided system can distinguish between multiple targets only when they are well separated. In this simulation we set the threshold between two separable targets to be 3°. For the purpose of evaluating the metrics, a measurement is considered to be a false positive if the distance between the measurements and its corresponding ground truth position is greater than 20 cm.

Figure 3.21 shows different scenarios which occur during multi-target tracking. At time t = 1, the 4 targets are well separated and can be identified uniquely. The hollow circles show the ground truth locations of the targets and the filled circles show the corresponding measurements. The colors indicate the correspondence between the ground truth positions of the targets and their measurements. A false positive is shown for target 1 at time t = 2 since the distance between the measurement and its corresponding ground truth is greater than 10 cm and the angular separation is greater than 3°. At time t=28, targets 1, 2 and 3 are close to each other causing an overlap. The ground truth location of target 2 is closest to the measurement and is therefore associated with it while targets 1 and 3 are considered missed. A switch in identities can be observed at time t = 31, with targets 1, 2 and 4 being labeled incorrectly.

Table 3.2 shows the performance of a multi-target tracking system and the



(a) t = 1 (b) t = 2

(c) t = 28 (d) t = 31

Figure 3.21: Tracking scenarios occuring in a multi-target tracking system.

relationship with an increase in the number of targets. From the table it can be observed that for a single target, the performance is highest as expected. An increase in the number of targets is associated with a decrease in performance. It should be noted that the performance of the system also depends on the interaction between targets. If tracks are simulated such that the targets do not come close to each other at any point in time, then the errors due to mismatches and misses will be 0% and the performance of the system will be close to 100% with errors only due to false positives.

# of subjects	# of instances	Misses	False positives	Mismatches	MOTA (%)	MOTP (%)
1	55	0	2	0	96.36	9.76
2	110	6	2	1	91.82	9.27
3	165	23	4	2	82.42	9.82
4	220	37	4	5	79.09	12.35

Table 3.2: Performance of simulated multi-target tracking system using LADAR.

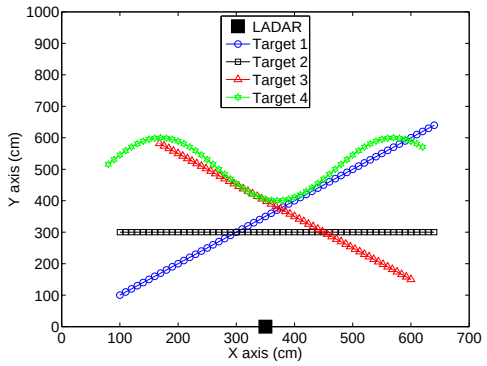


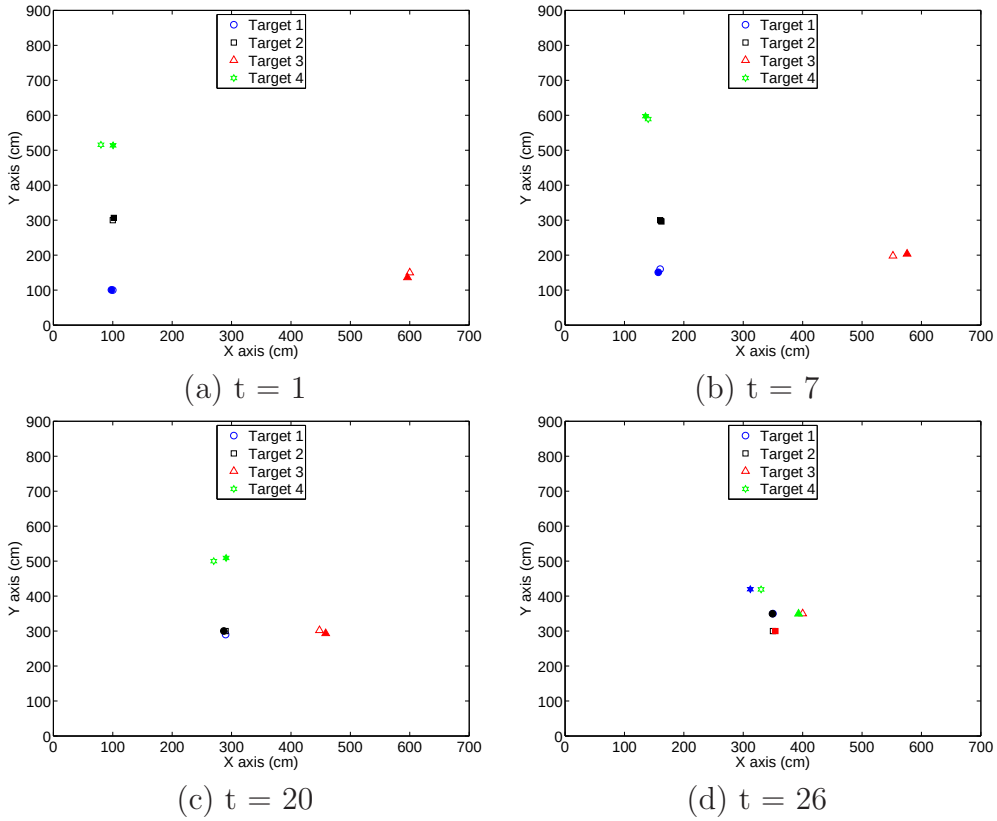
Figure 3.22: Simulated tracks for camera based tracking system.

3.2.7.2 Camera

This simulation demonstrates the MOT metrics on a 2D camera based tracking system. A room of similar dimensions, $700 \times 700\ cm^2$ is considered. Here we consider a network of cameras placed around the room. Using the cameras an occupancy map is created to extract the targets from the background. The targets are represented as blobs with their x, y position given by the centroids of the blobs. As in the previous simulation we assume that there are 55 measurements for each object available at equal time intervals. The ground truth positions of the tracks are shown in Figure 3.22. We have added a Gaussian noise with a standard deviation of 10 cm to the ground truth positions to create measurements. In this simulation we assume that when the Euclidean distance between two targets is less than 30 cm, they overlap each other and cannot be distinguished. Another assumption is that if the Euclidean distance between the hypothesis and the target is greater than 20 cm then it is

considered a false positive.

Figure 3.23 shows different cases where correspondence errors occur in this simulation. At time t = 1, all measurements are well separated and associated with a unique target. The measurement associated with target 3 at time t = 2 is greater than 20 cm and is considered a false positive. At time t = 20, targets 1 and 3 overlap each other since they are closer than 30 cm away from each other. Hence, the measurement is associated only with target 1. This is considered as a missed error for target 3. We can observe switching of identities at time t = 26. This is an example of mismatched error.



(a) t = 1 (b) t = 7

(c) t = 20 (d) t = 26

Figure 3.23: Tracking scenarios occuring in a multi-target tracking system.

Table 3.3 shows the performance of the simulated multi-target tracking system with increasing number of targets. From the figure it can be observed that the miss

and mismatch errors are 0 as expected. As the number of targets increases the probability of mismatches and misses increases as the targets overlap each other. This can seen with the corresponding decrease in the MOTA.

# of subjects	# of instances	Misses	False positives	Mismatches	MOTA (%)	MOTP (%)
1	55	0	5	0	90.91	9.35
2	110	5	6	2	88.18	8.43
3	165	8	19	6	80.00	9.47
4	220	19	20	27	70.00	11.16

Table 3.3: Performance of simulated multi-target tracking system using camera.

3.3 Experimental Results

3.3.1 UWB and LADAR

Figures 3.24 – 3.31 show the results as the number of people in the room is increased from 1 to 8. In each of these figures, plots (a)–(c) show the Euclidean error over time, while plots (d)–(f) show the x,y filtered output. Figure 3.32 shows the average increase in the data association metrics per person versus the number of people in the room, while Figure 3.33 shows the MOTA and MOTP curves for the LADAR and UWB sensors and their fusion.

From the figures we can observe that when there is only one person in the room, the LADAR measurements are not occluded and its precision is better than that of the UWB. With an increase in the number of people from 1 to 8, the instances of occlusions for the LADAR sensor increases leading to a rise in missed measurements. In addition, the number of false positives and mismatches also increase due to incorrect data association (shown in Figure 3.32 (a)). The increase in error due to missing measurements in the LADAR can also be observed from the plots 3.24 – 3.31 (a). The rise in errors leads to a decrease in MOTA and is reflected in Figure

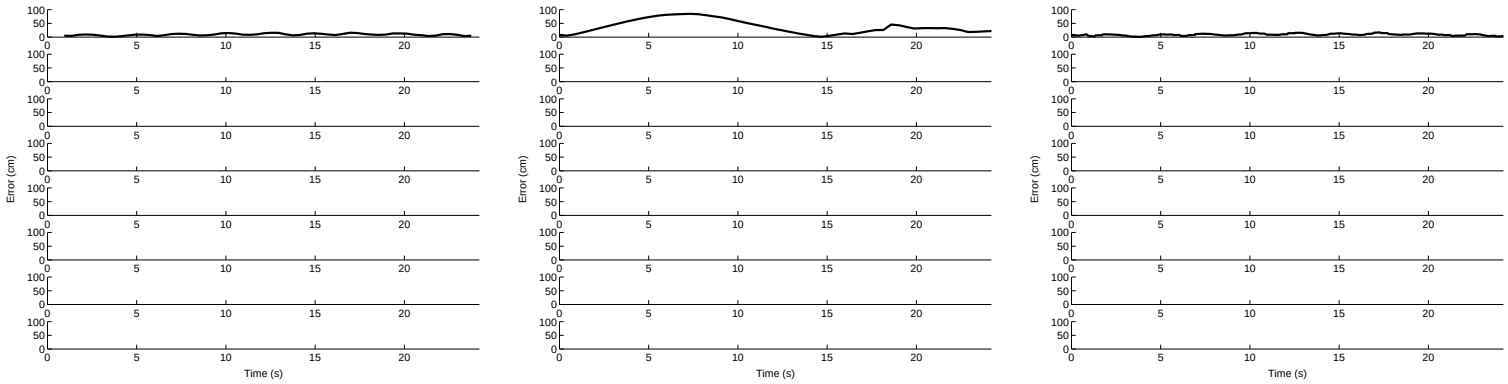
3.33 (a). From the figure we can observe that the MOTA of the LADAR decreases from 100% to approximately 67%. As the number of people in the room increases from 1 to 8, the precision of the LADAR also worsens from approximately 17 cm to 34 cm (Figure 3.33 (b)). On the other hand, the accuracy and precision of the UWB remains relatively constant, around 95% and 31 cm respectively even with a change in the number of people in the room.

The MOTA of the fusion is better than either the LADAR or UWB sensor alone, and is close to 100%. This shows that the UWB measurements can help solve the data association problem for the LADAR sensor. From Figure 3.33 (b) we can observe that for 1 person in the room, the fusion improves the precision of tracking by approximately 40% as compared to that of the UWB sensor. The MOTP curve for the fusion follows the LADAR up to 6 people, improving the precision of tracking over the UWB sensor alone. This is because the LADAR sensor is assumed to have a better noise model with a lower standard deviation as compared to the UWB. Therefore, the filter tends to weight the LADAR measurements more than those from the UWB sensor. When the error in the LADAR is more than the UWB, the fused output tends to be closer to that of the UWB since the LADAR is no longer able to provide precise measurements.

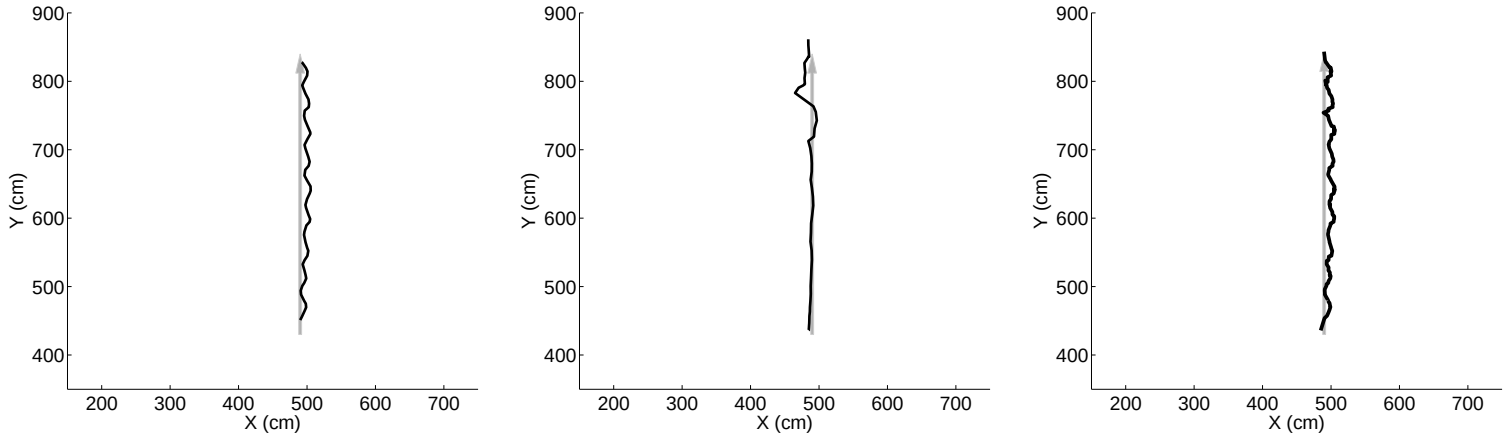
3.3.2 UWB and Camera

As in the previous section, Figures 3.34 – 3.41 show the results as the number of people in the room is increased from 1 to 8. Figures 3.42 and 3.43 show the accuracy metrics and MOT curves for the camera and UWB sensors and their fusion.

As in the case of LADAR sensor, we can observe that when there is only one person in the room, the accuracy of the camera is 100%. When the number of people

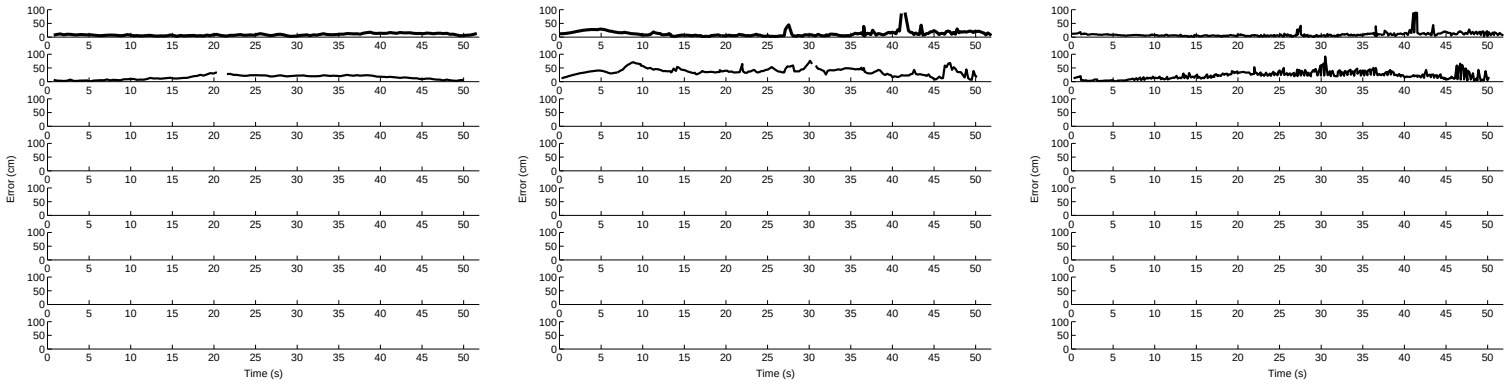


(a) LADAR (b) UWB (c) LADAR+UWB

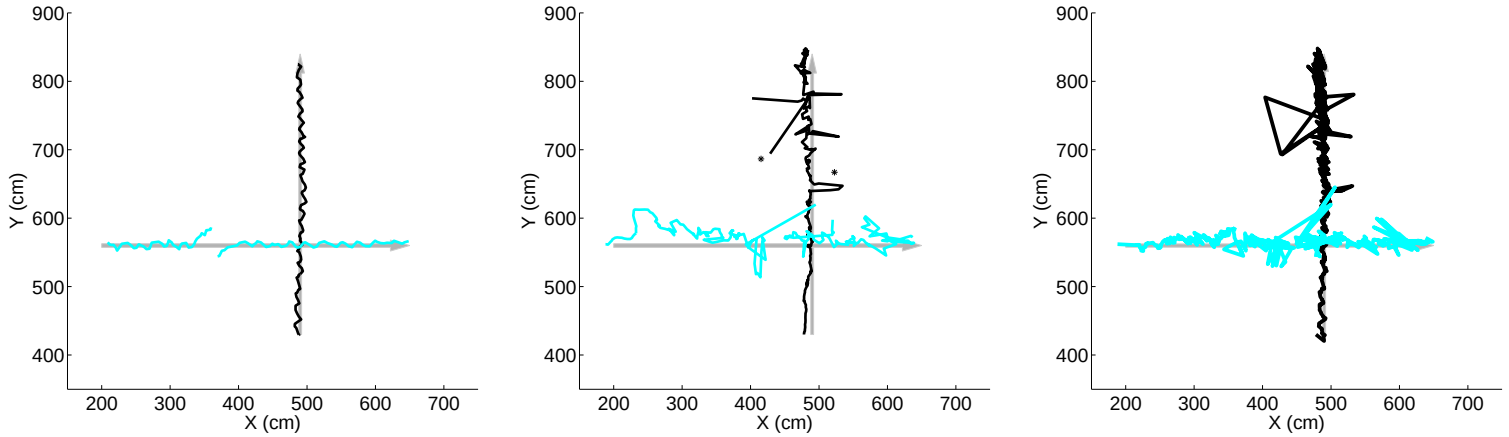


(d) LADAR (e) UWB (f) LADAR+UWB

Figure 3.24: Multi-object tracking with 1 people in the room.

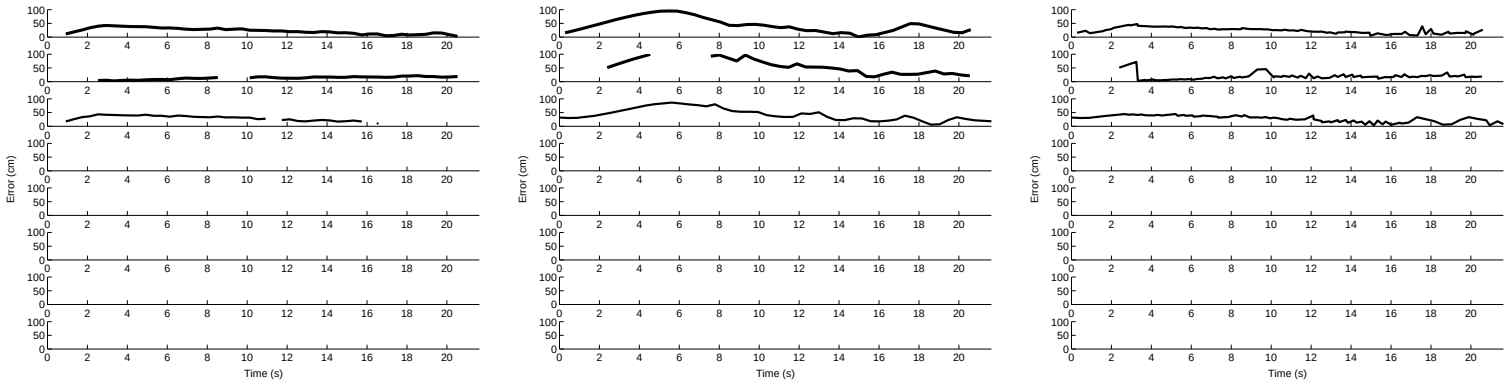


(a) LADAR (b) UWB (c) LADAR+UWB

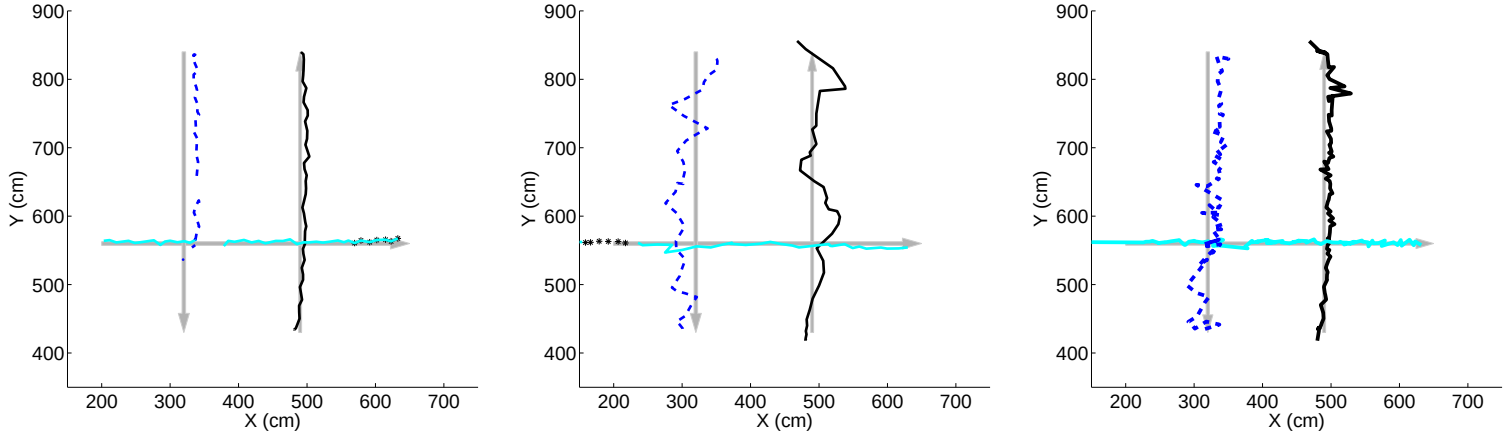


(d) LADAR (e) UWB (f) LADAR+UWB

Figure 3.25: Multi-object tracking with 2 people in the room.

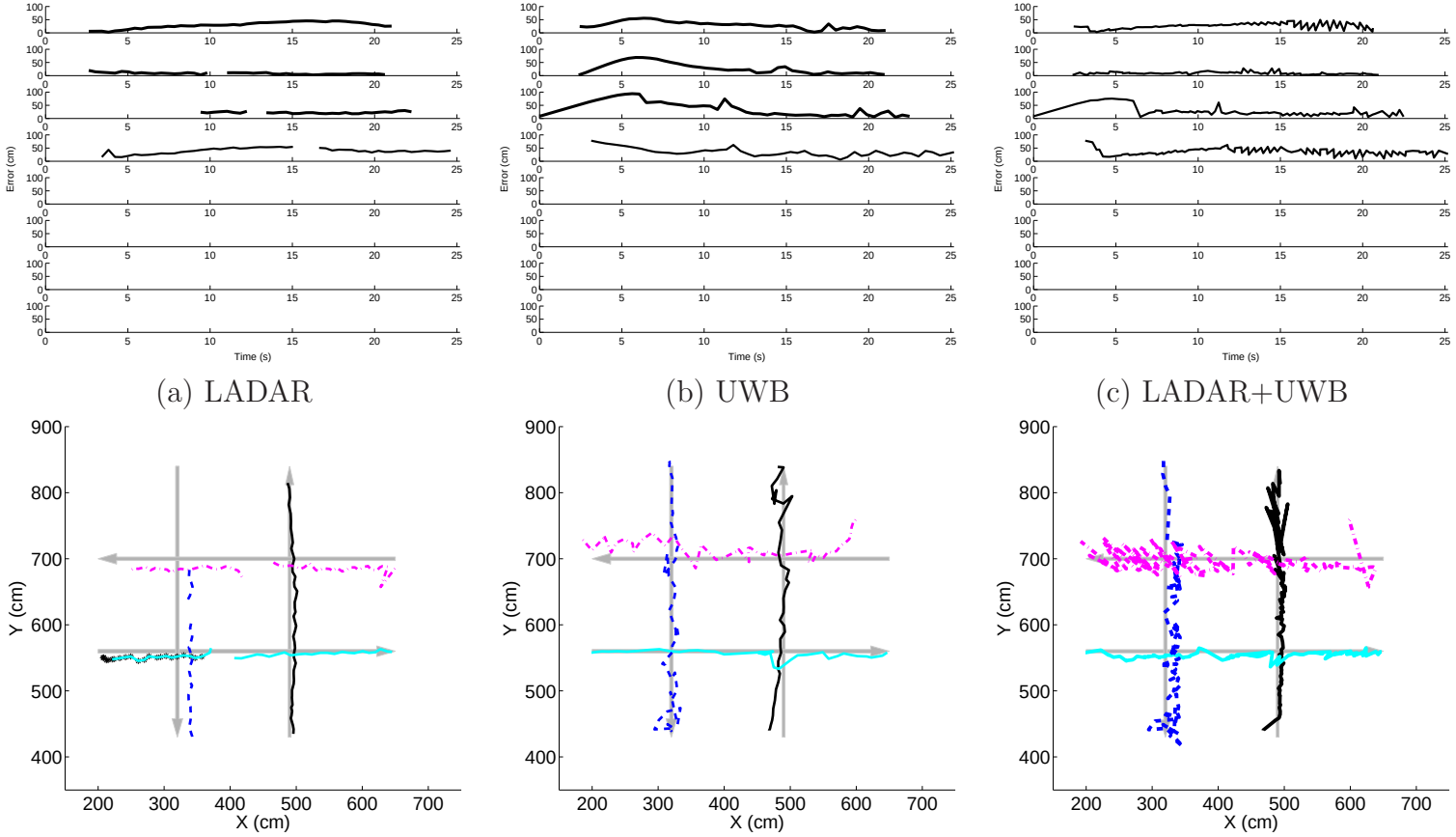


(a) LADAR (b) UWB (c) LADAR+UWB



(d) LADAR (e) UWB (f) LADAR+UWB

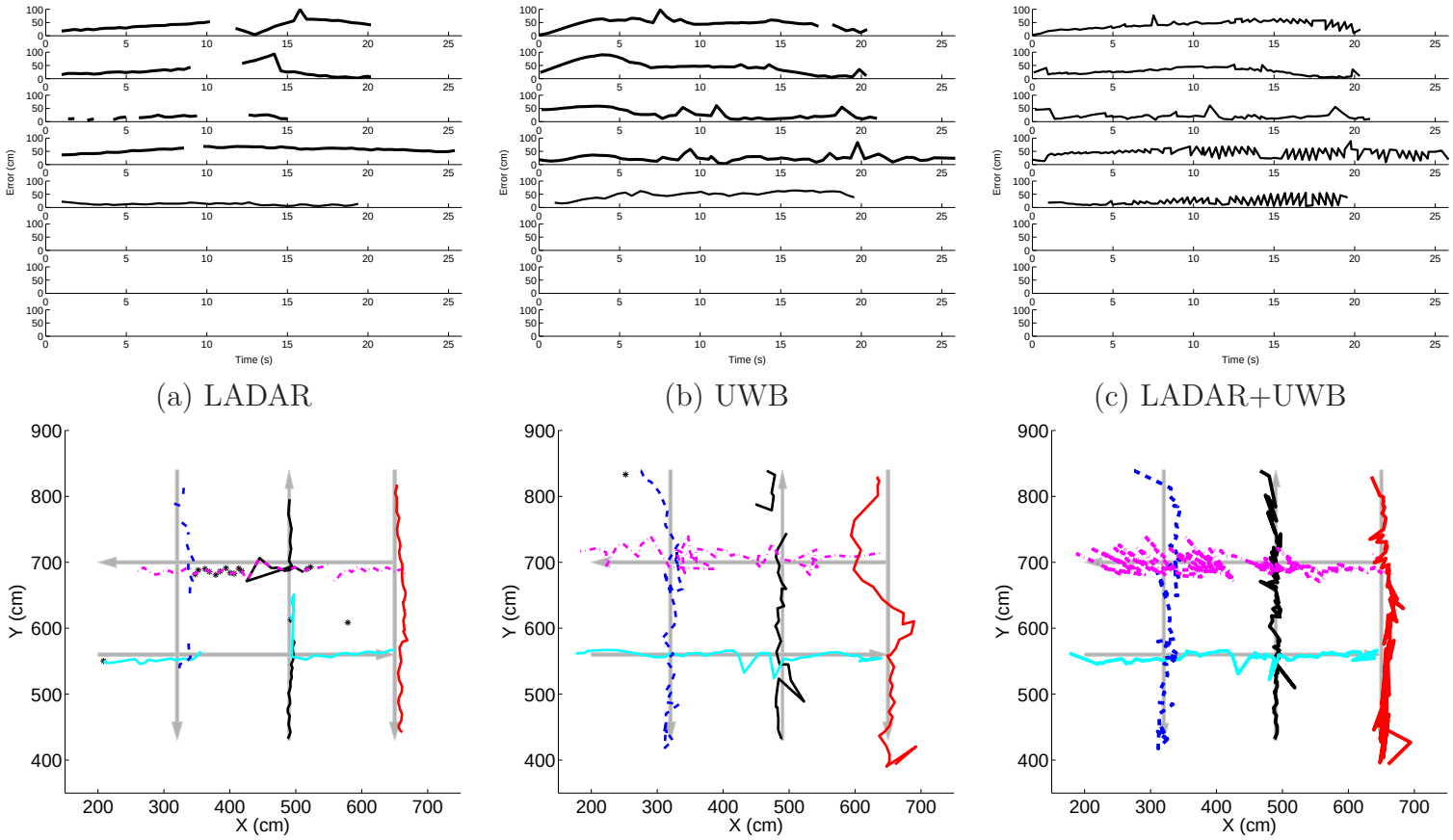
Figure 3.26: Multi-object tracking with 3 people in the room.



(a) LADAR (b) UWB (c) LADAR+UWB

(d) LADAR (e) UWB (f) LADAR+UWB

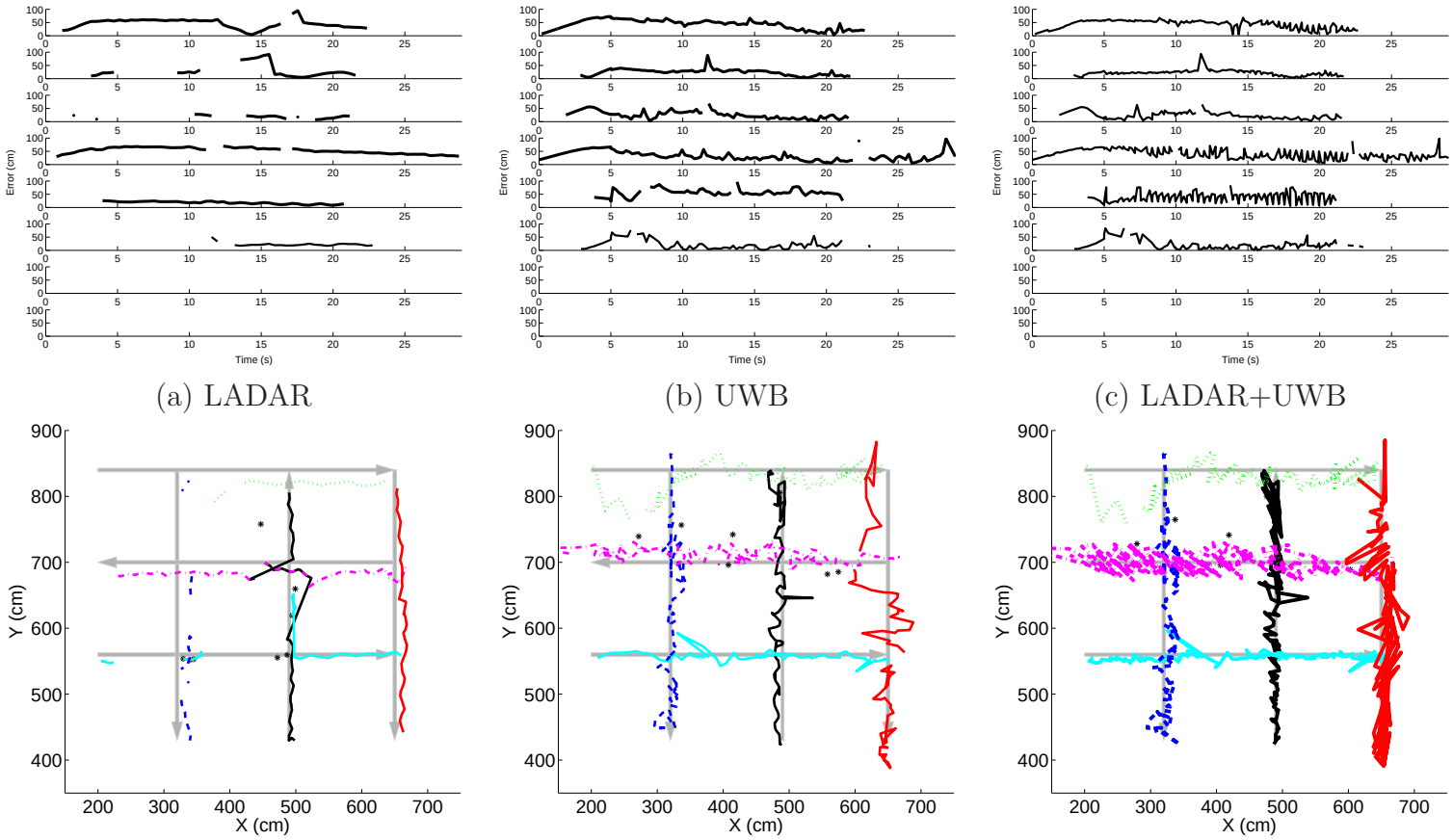
Figure 3.27: Multi-object tracking with 4 people in the room.



(a) LADAR (b) UWB (c) LADAR+UWB

(d) LADAR (e) UWB (f) LADAR+UWB

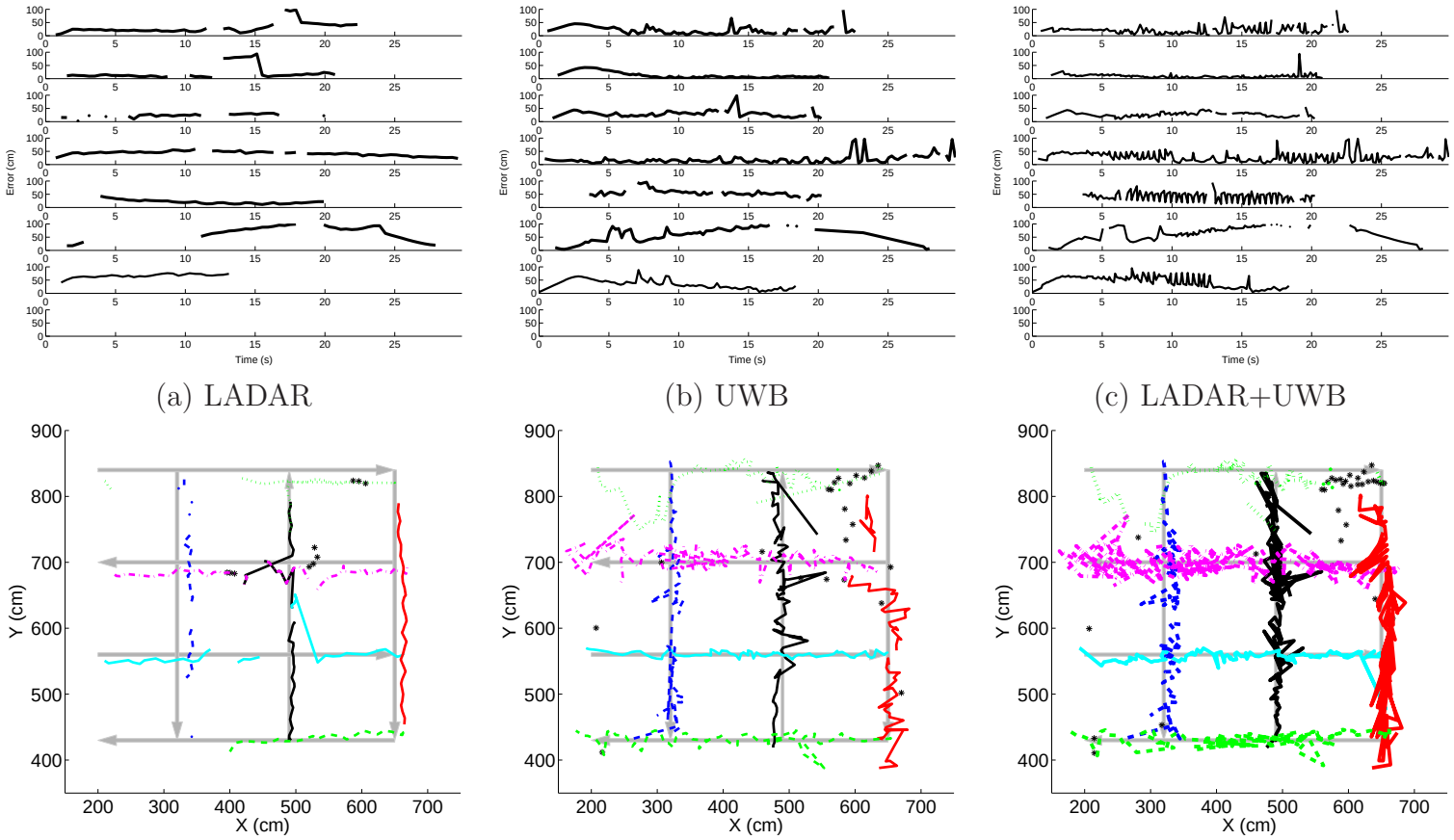
Figure 3.28: Multi-object tracking with 5 people in the room.



(a) LADAR (b) UWB (c) LADAR+UWB

(d) LADAR (e) UWB (f) LADAR+UWB

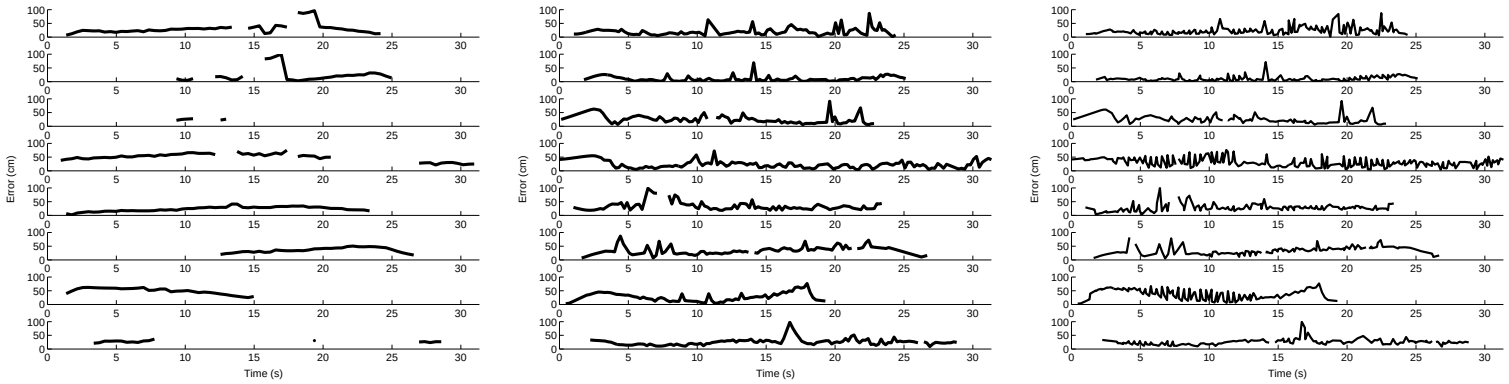
Figure 3.29: Multi-object tracking with 6 people in the room.



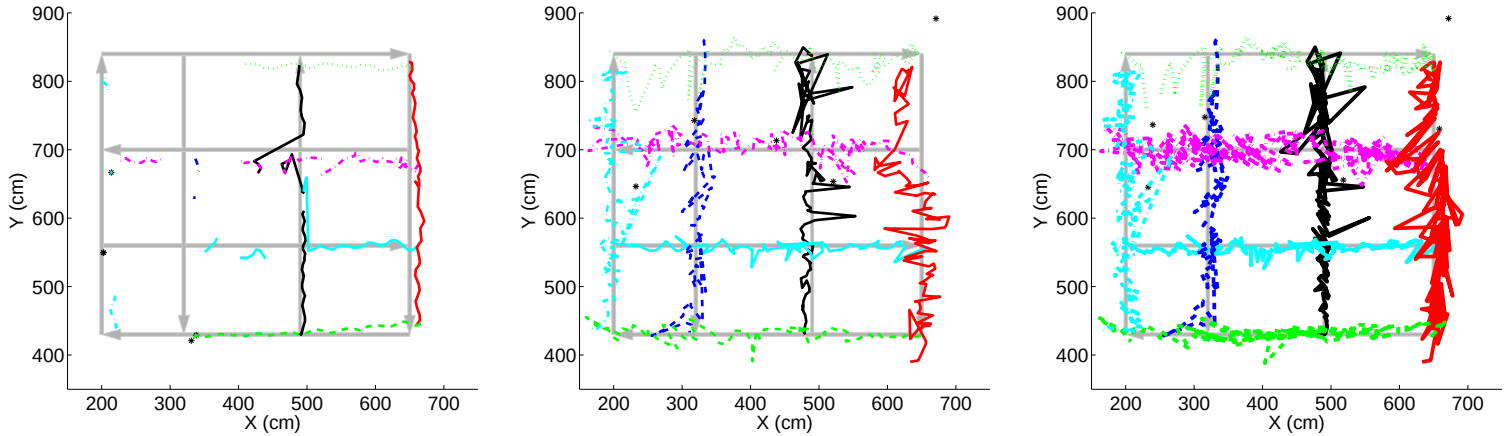
(a) LADAR (b) UWB (c) LADAR+UWB

(d) LADAR (e) UWB (f) LADAR+UWB

Figure 3.30: Multi-object tracking with 7 people in the room.

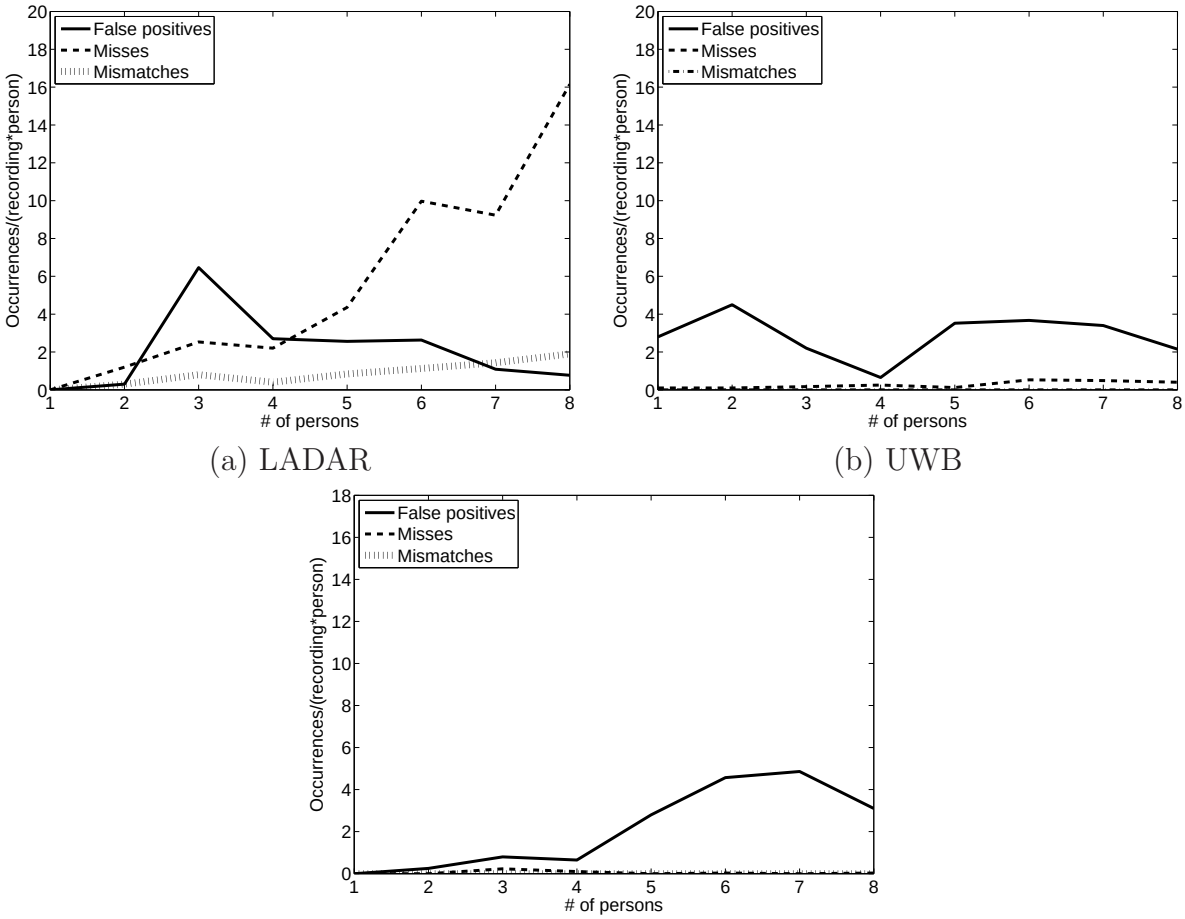


(a) LADAR (b) UWB (c) LADAR+UWB



(d) LADAR (e) UWB (f) LADAR+UWB

Figure 3.31: Multi-object tracking with 8 people in the room.

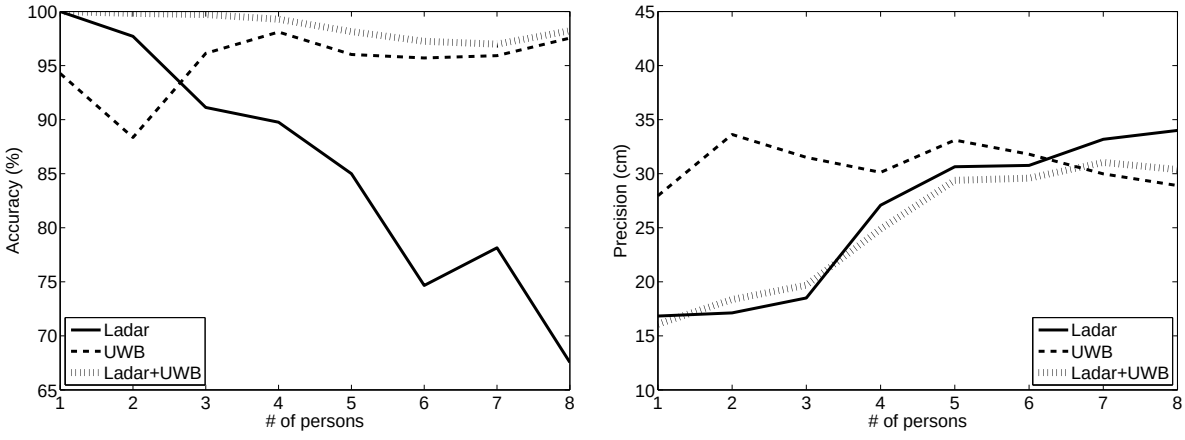


(a) LADAR

(b) UWB

(c) LADAR and UWB

Figure 3.32: Accuracy metrics for LADAR and UWB sensors.



(a) MOTA

(b) MOTP

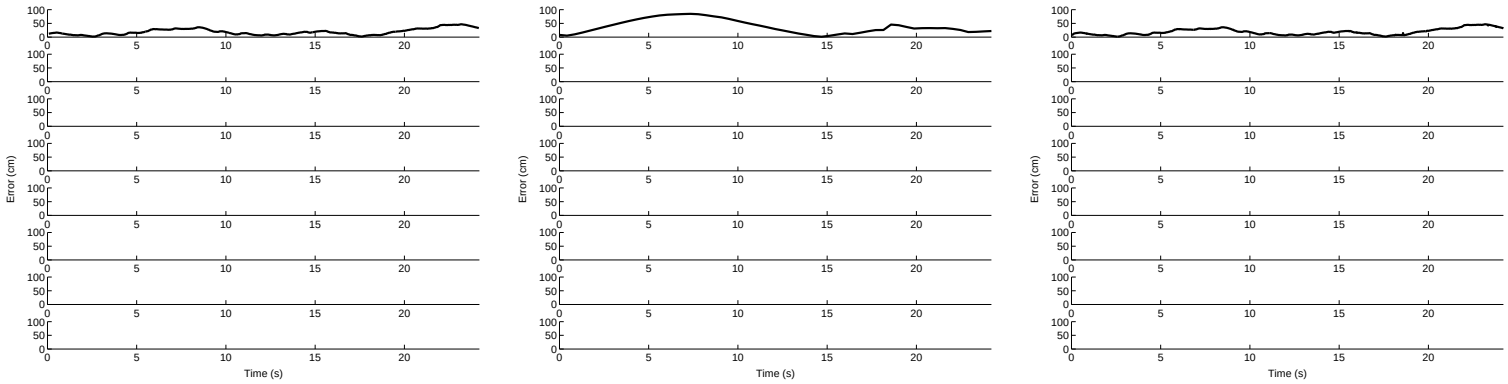
Figure 3.33: MOTA and MOTP for LADAR and UWB sensors.

in the room increases, the number of distinct blobs present decreases. This results in missed, mismatch and false positive errors which in turn decreases the MOTA. A specific example of errors caused by missing measurements can be seen in the Figure 3.41 (a). The increase in errors is reflected in a decrease in the MOTA for the camera from 100% for up to 2 people to 98% for 8 people. The increase in the number of people also corresponds with a decreasing precision due to increase in the number of data association errors. The MOTA for the camera does not drop as dramatically as observed in the case of LADAR as shown in Figure 3.43 (a). This is due to the presence of the large number of measurements available for the camera due to its higher speed of operation ($\approx$ 20 Hz). From Figure 3.43 (b) it can be observed that the MOTP of the camera worsens from 23 cm for 1 person to 37 cm for 8 people. As described in the previous section, the MOTA and MOTP of the UWB remains relatively constant at 95% and 31 cm respectively.

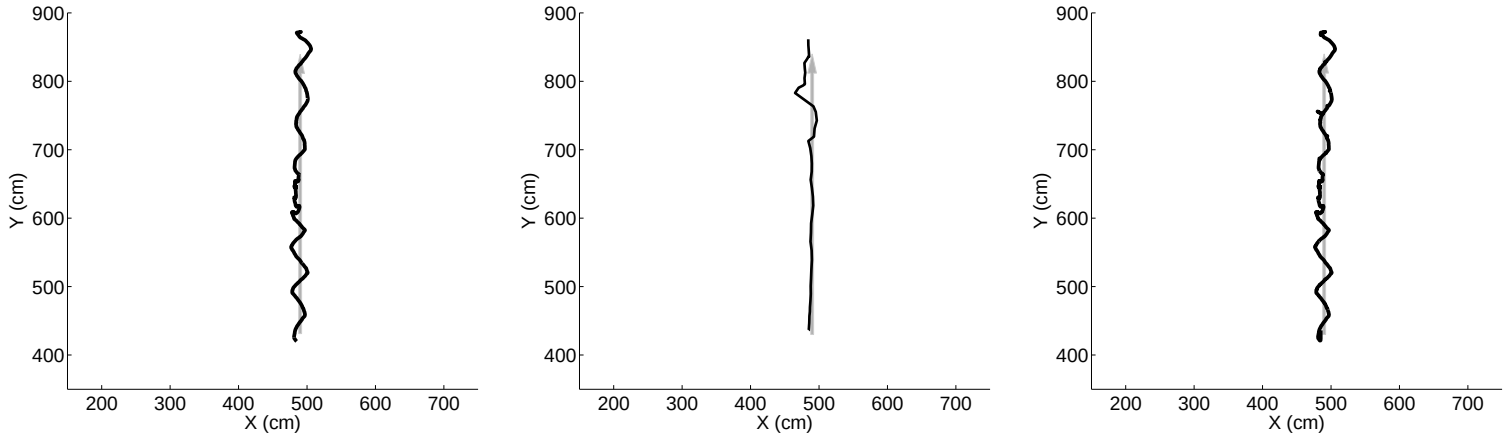
The fusion improves the MOTA for 1 to 8 people with an accuracy around 98%. However, the MOTP curve follows the same trend as the camera; it improves the precision up to 2 people and then the precision worsens from 22 to 32 cm. The fusion results are closer to that of the camera due to the larger number of measurements present as compared to that of the UWB (roughly 4 to 1).

3.4 Conclusions

The work in this chapter was motivated by the idea of a differential GPS. We explored whether the same idea could be implemented for a UWB based LPS. The objective was to improve the precision of the UWB sensor by augmenting it with a differential sensor placed in each room. In our experiments we tested using a LADAR and camera network as differential sensors.

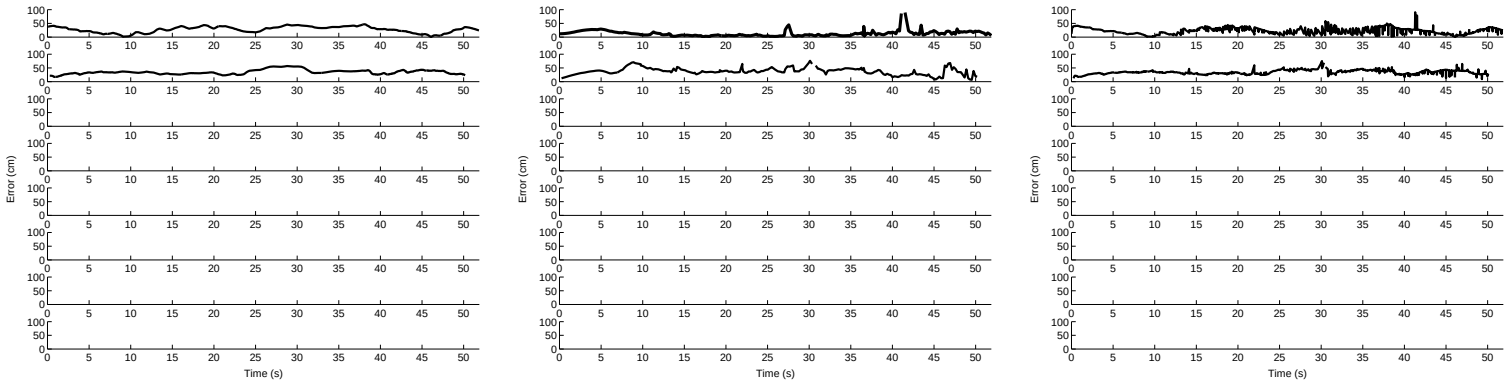


(a) Camera (b) UWB (c) Camera+UWB

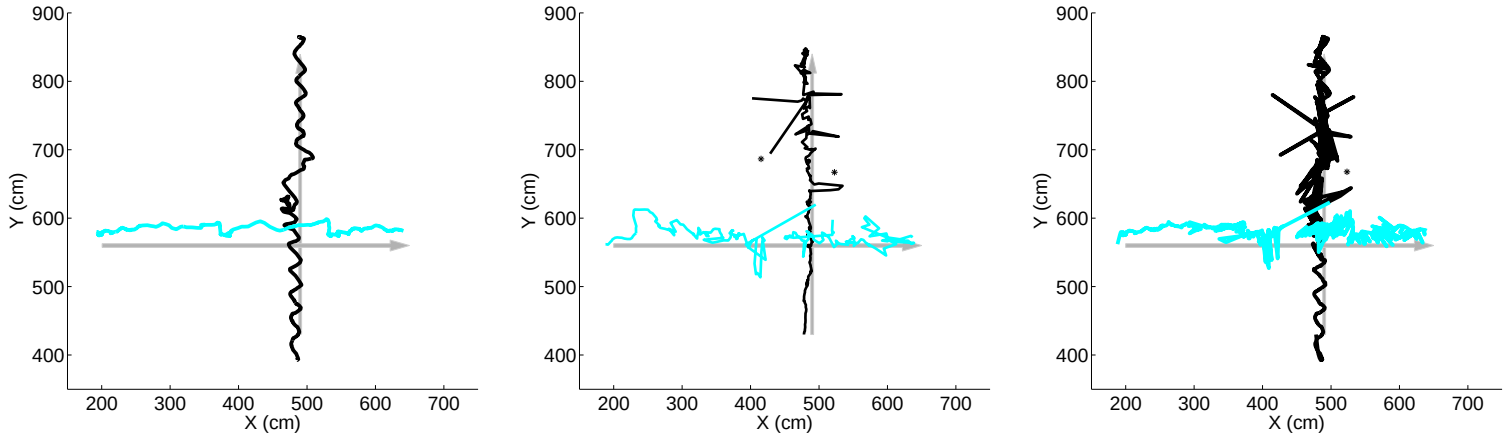


(d) Camera (e) UWB (f) Camera+UWB

Figure 3.34: Multi-object tracking with 1 person in the room.

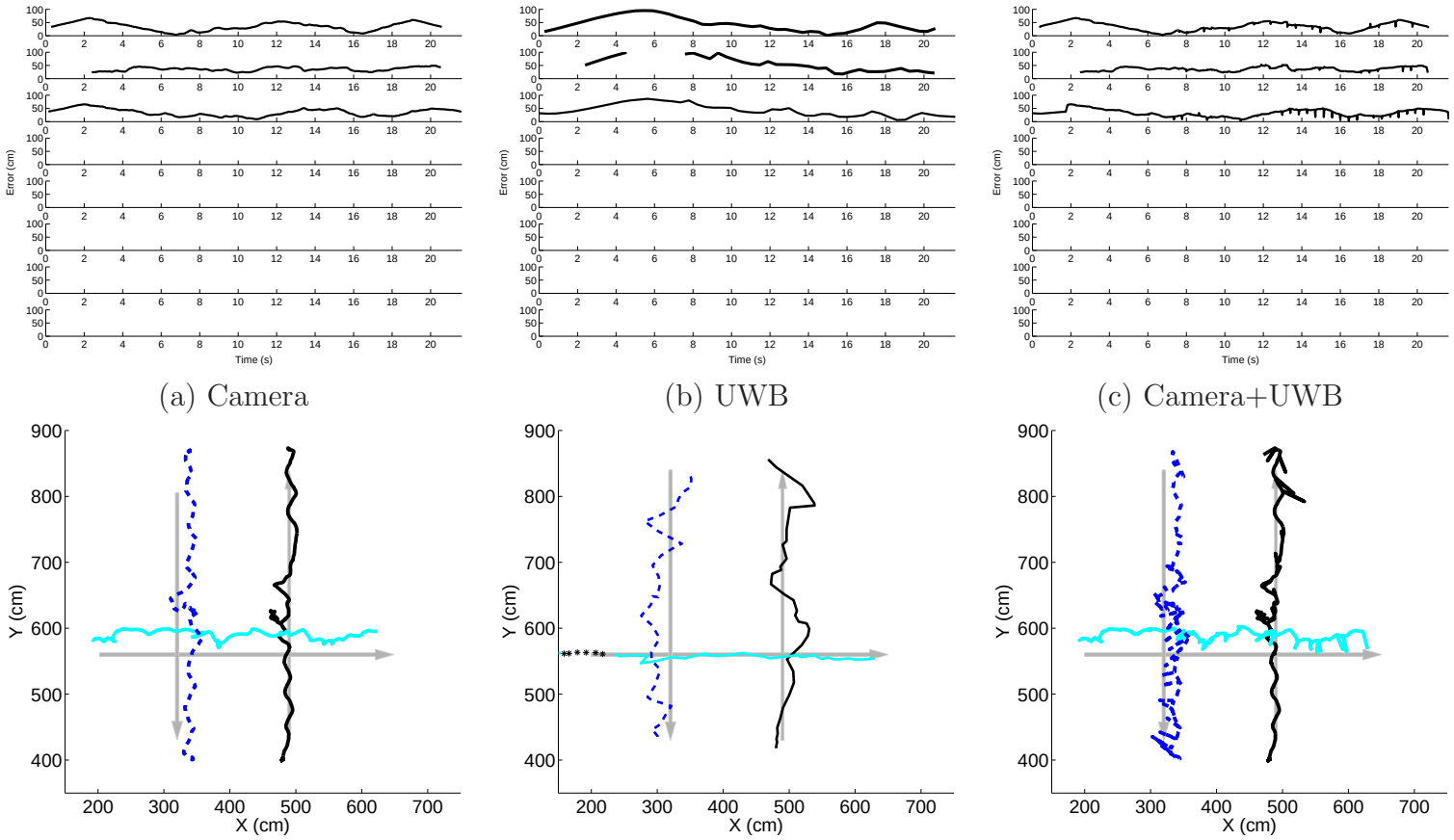


(a) Camera (b) UWB (c) Camera+UWB



(d) Camera (e) UWB (f) Camera+UWB

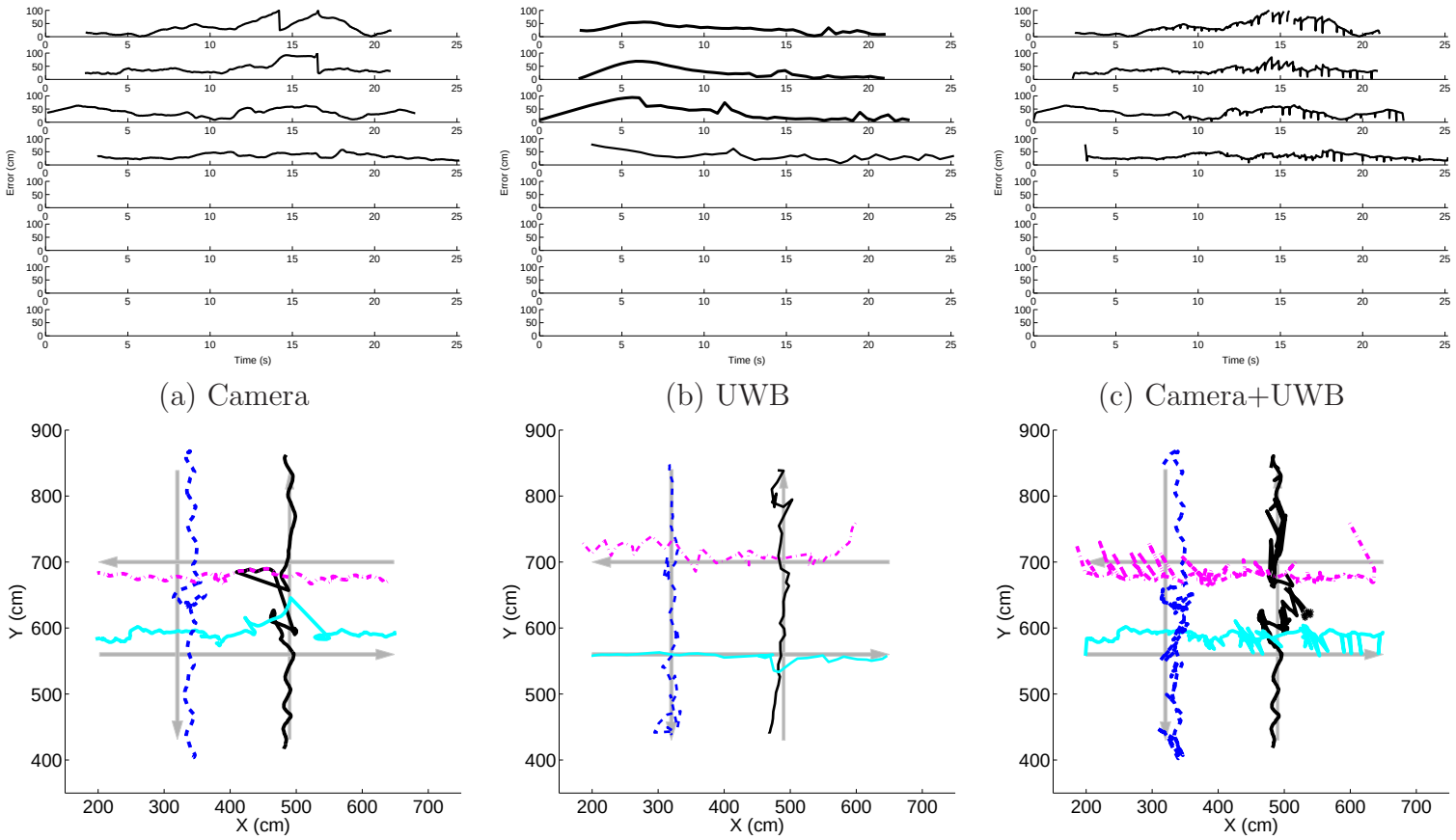
Figure 3.35: Multi-object tracking with 2 people in the room.



(a) Camera (b) UWB (c) Camera+UWB

(d) Camera (e) UWB (f) Camera+UWB

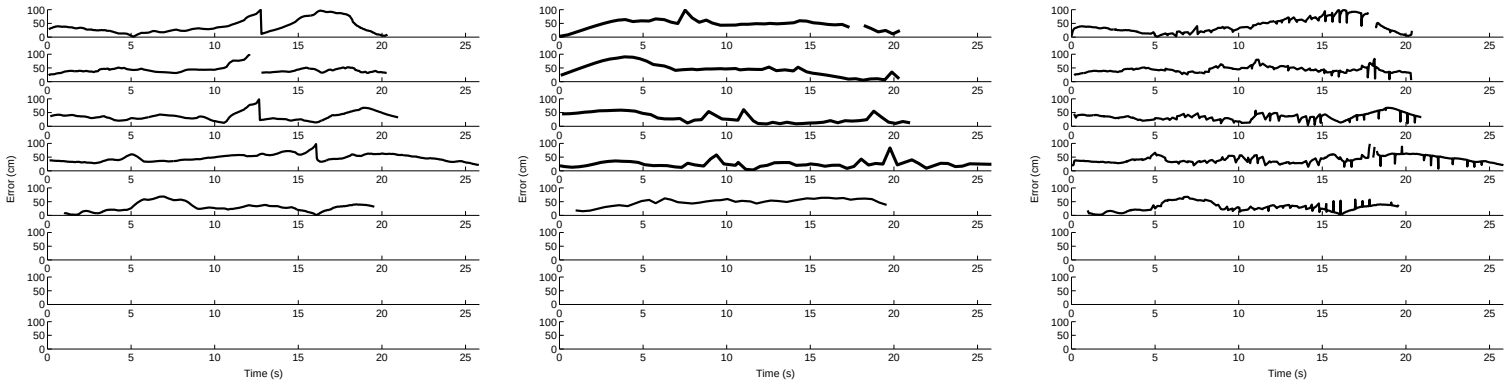
Figure 3.36: Multi-object tracking with 3 people in the room.



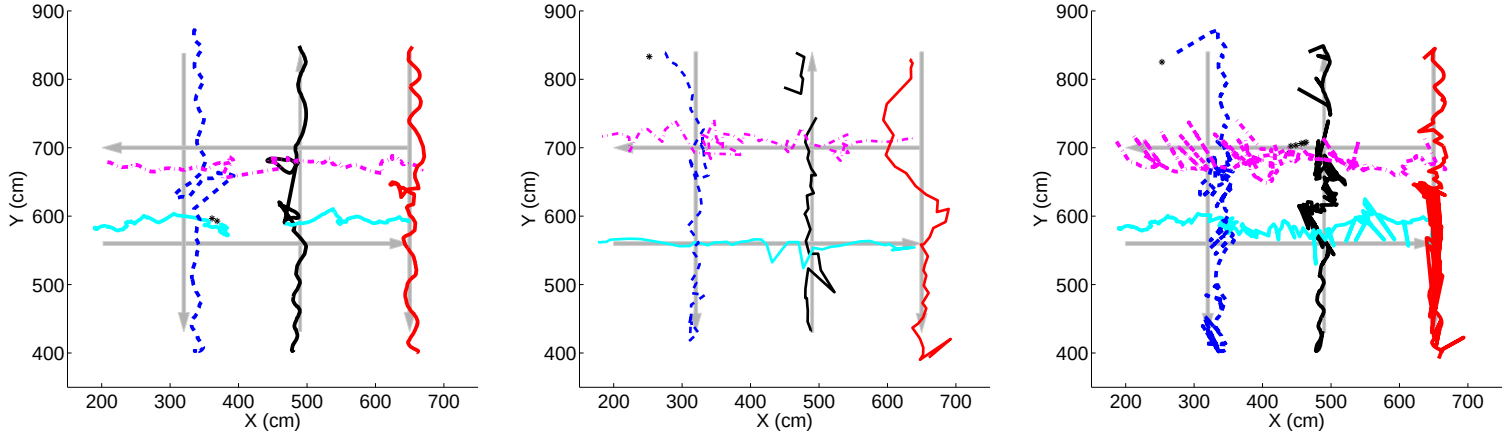
(a) Camera (b) UWB (c) Camera+UWB

(d) Camera (e) UWB (f) Camera+UWB

Figure 3.37: Multi-object tracking with 4 people in the room.

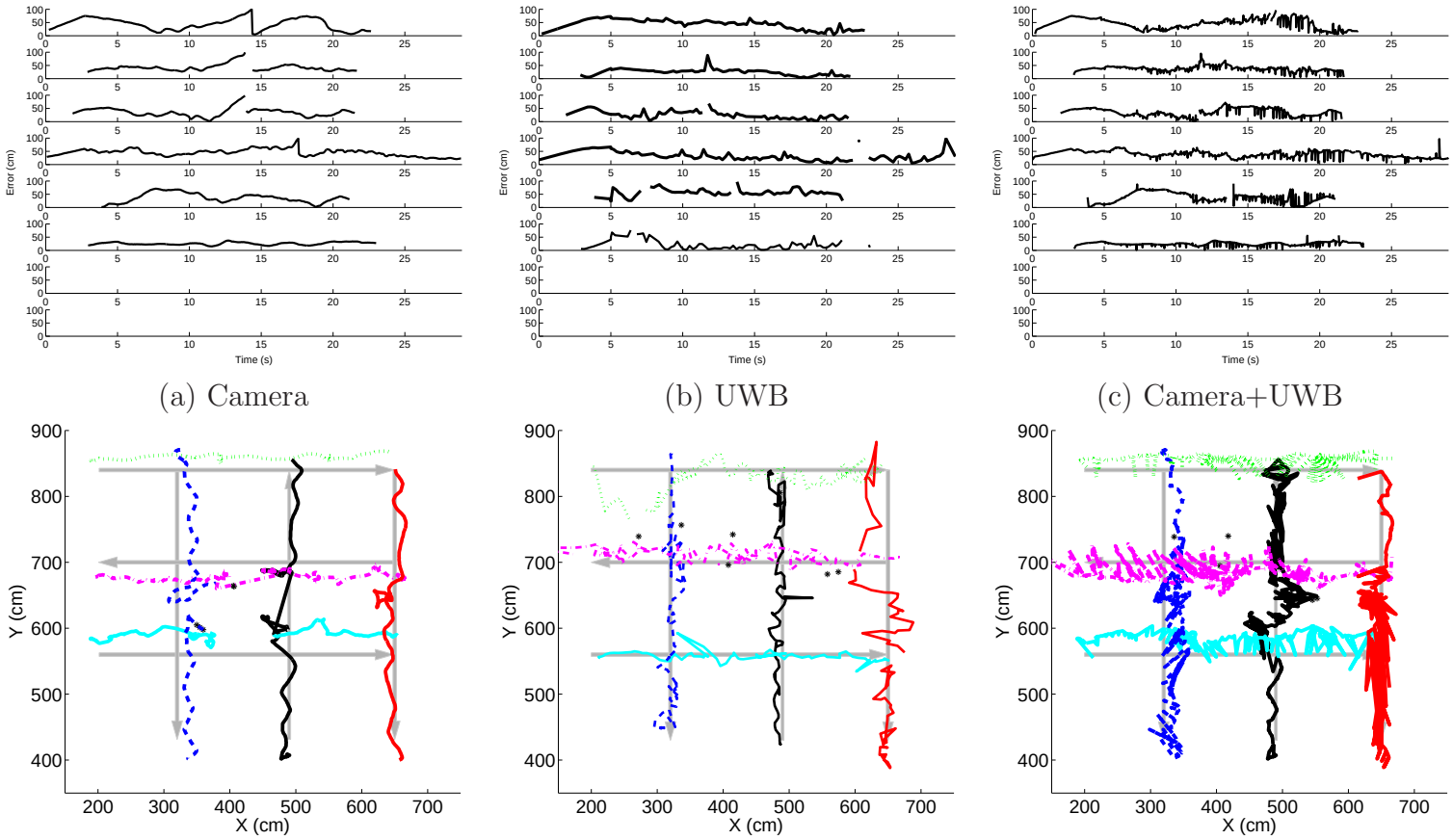


(a) Camera (b) UWB (c) Camera+UWB



(d) Camera (e) UWB (f) Camera+UWB

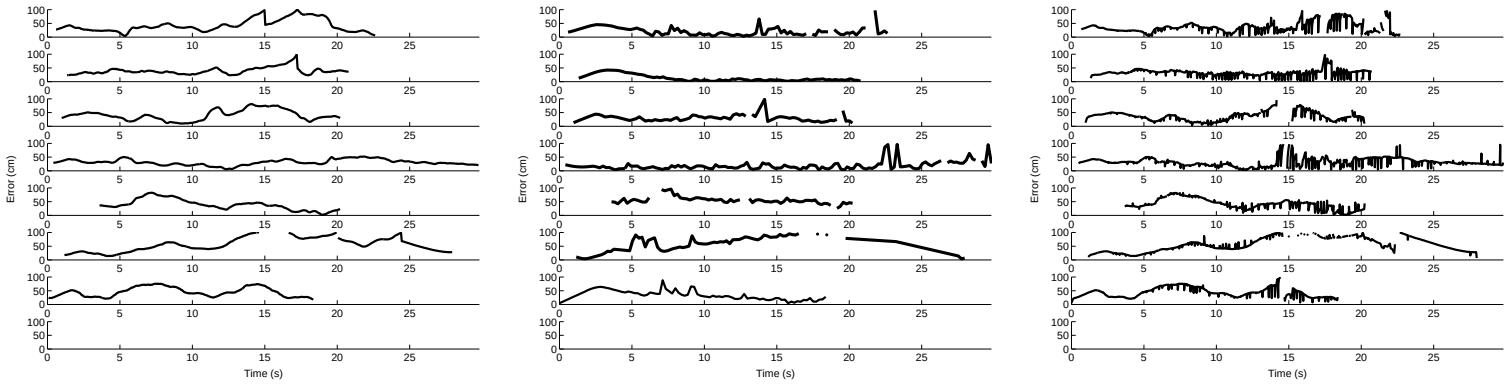
Figure 3.38: Multi-object tracking with 5 people in the room.



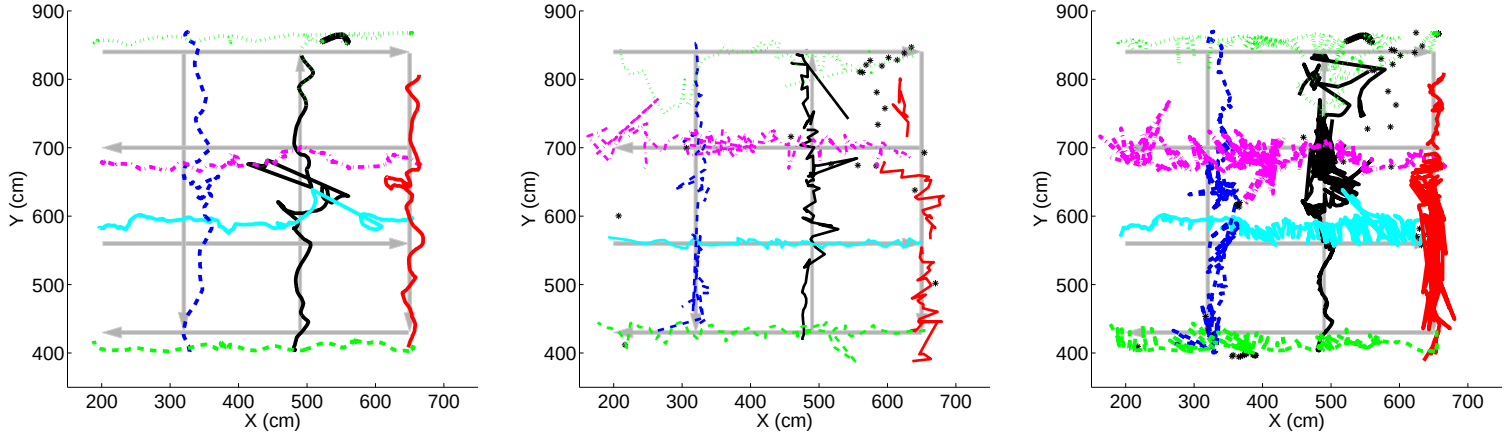
(a) Camera (b) UWB (c) Camera+UWB

(d) Camera (e) UWB (f) Camera+UWB

Figure 3.39: Multi-object tracking with 6 people in the room.

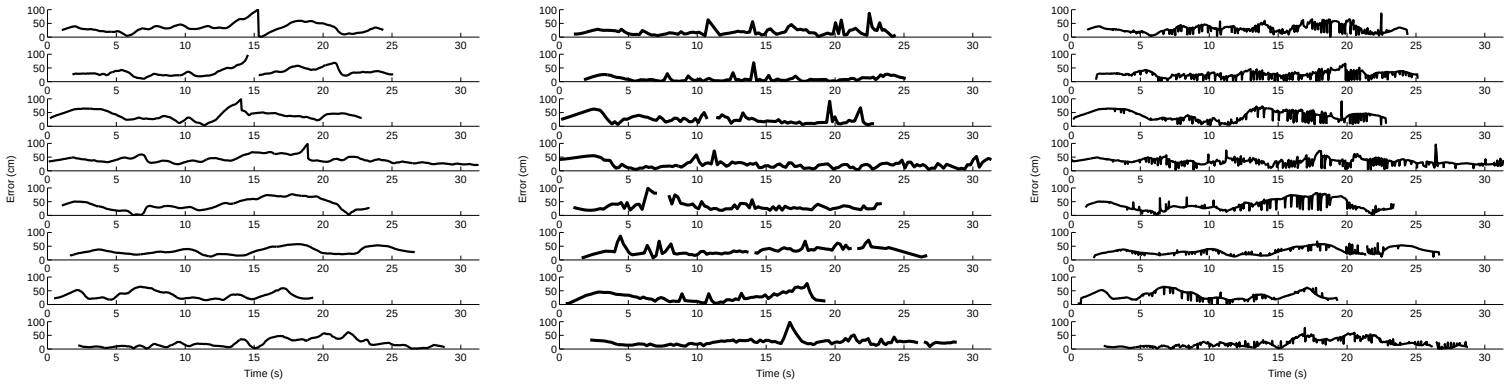


(a) Camera (b) UWB (c) Camera+UWB

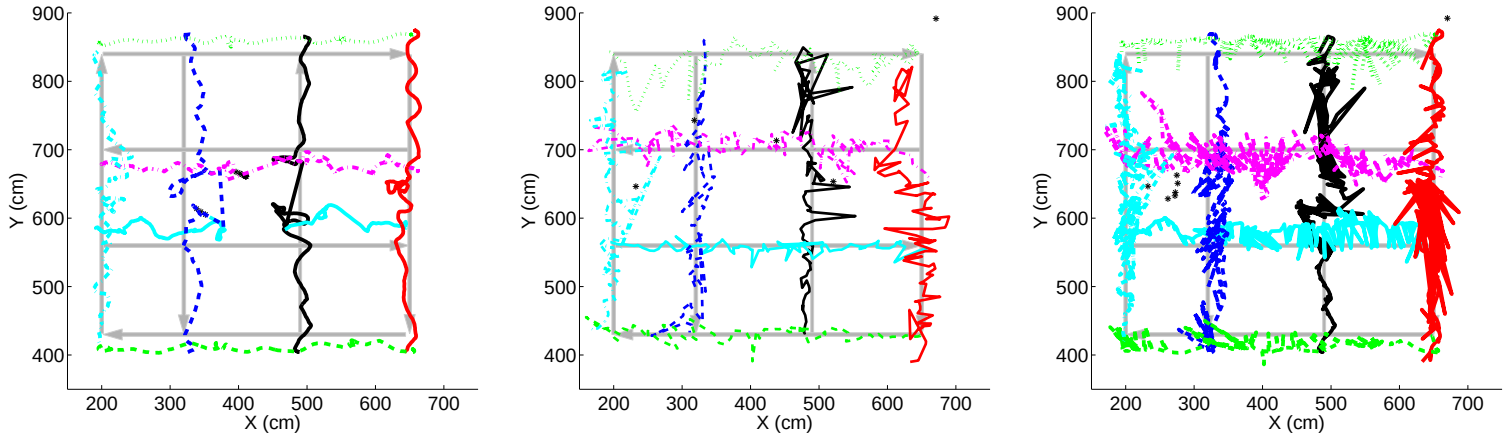


(d) Camera (e) UWB (f) Camera+UWB

Figure 3.40: Multi-object tracking with 7 people in the room.

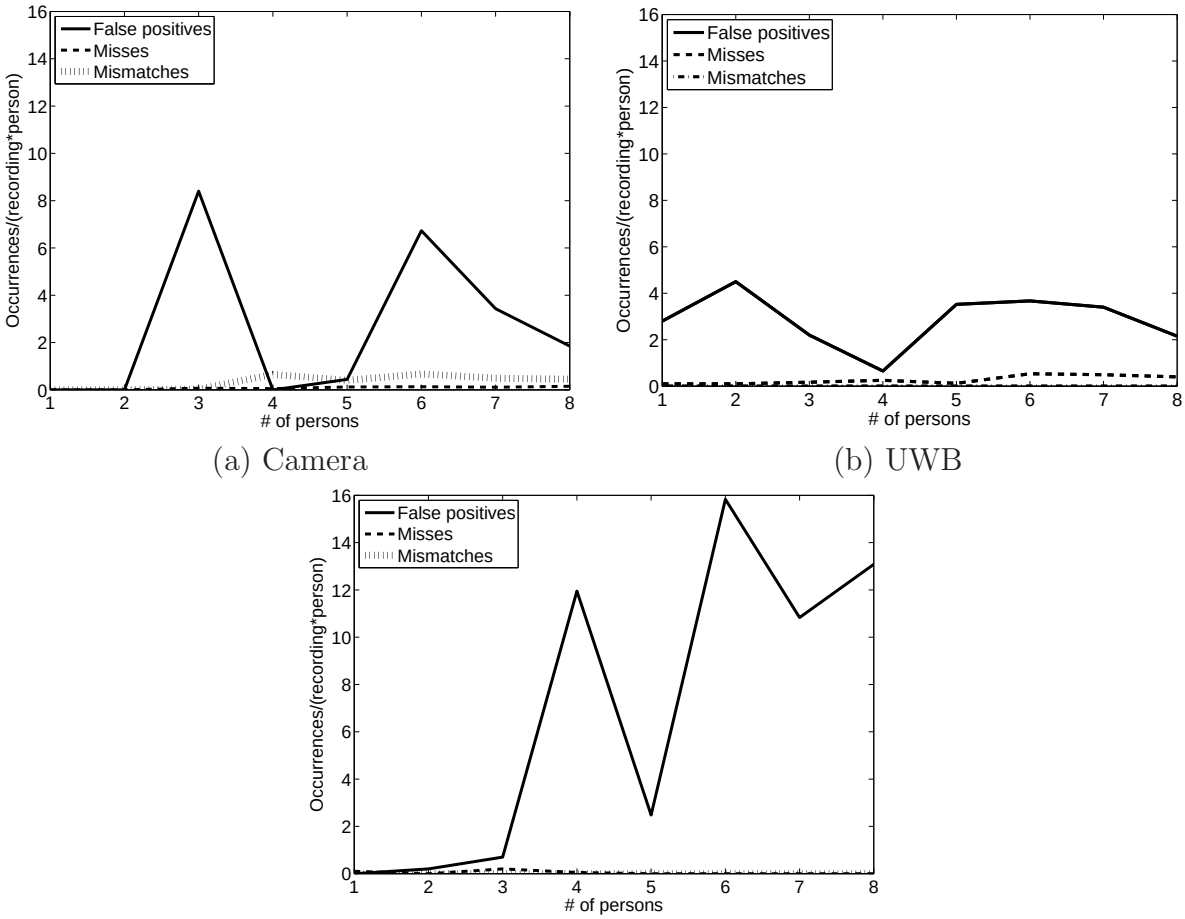


(a) Camera (b) UWB (c) Camera+UWB



(d) Camera (e) UWB (f) Camera+UWB

Figure 3.41: Multi-object tracking with 8 people in the room.

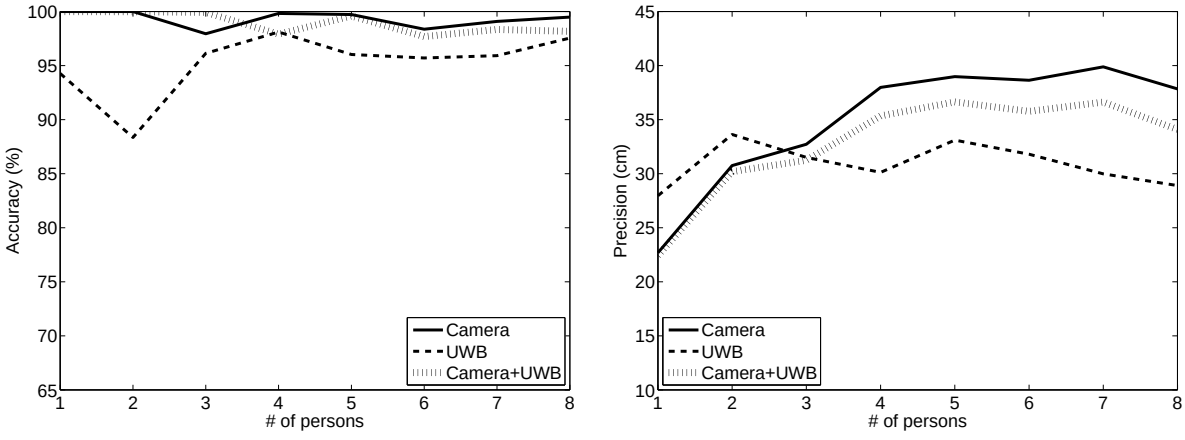


(a) Camera

(b) UWB

(c) Camera and UWB

Figure 3.42: Accuracy metrics for camera and UWB sensors.



(a) MOTA

(b) MOTP

Figure 3.43: MOTA and MOTP for camera and UWB sensors.

From our results, we observed that the fusion of the UWB and LADAR sensors was able to improve precision in tracking for up to 6 people. The precision of the UWB was almost constant around 31 cm. On the other hand the precision of the LADAR was roughly 17 cm for 1 person and became worse as the number of people in the room increased. The fusion improved the precision by roughly 40% when tracking 1 person as compared to that of the UWB. Even with three people in the room the fused output was able to improve precision by approximately 25%. However, as the number of people increased beyond 6 the precision of the fusion between UWB and LADAR also became worse and was closer to that of the UWB. The MOTA for the UWB was also constant with an accuracy near 95%. The MOTA for the LADAR was 100% for 1 person tracking and decreased linearly as the number of people in the room increased. On the other hand, the fused output improved the accuracy for tracking 1-8 people. The fusion of the UWB and camera was able to improve precision for only up to 3 people and provide a modest improvement in precision, roughly 15% when there is only 1 person in the room. The fusion improved the MOTA and is close to 100%.

We used a Kalman filtering approach since we assumed that the noises in the sensor observations are zero-mean Gaussian. Our framework could easily be modified to use another type of filter and can be extended to other noise models. The expectation while developing the fusion framework was that using an aided (UWB) and unaided (LADAR or camera network) sensors would help solve the data association problem in the unaided sensor, while the unaided sensor would help inject higher precision measurements. Our experiments demonstrated the potential for this methodology, but somewhat failed to achieve the goal due to the low precision of the differential sensor. If the differential sensor did in fact provide a higher precision across the range of people tested, we expect that the fusion framework would have

resulted in higher precision than raw UWB measurements across all conditions. It would be useful to investigate the use of other sensors for differential measurements, such as RFID or Bluetooth.

Sensors	**# people**		
	1-2	**3-6**	**7-8**
UWB	≈ 31	≈ 31	≈ 31
LADAR	≈ 17	20–30	30–35
Camera	23–27	30–37	≈ 37
UWB+LADAR	≈ 16	17–27	≈ 30
UWB+Camera	22–27	30–32	≈ 32

Table 3.4: MOTP comparison (cm).

Table 3.4 compares the MOTP of the sensors and sensor combinations we tested as the number of people in the room increase. From the table it can be seen that when there are only 1-2 people in the room, the LADAR performs better than the UWB by approximately 40%. In this case the LADAR augments the UWB and improves its precision. When there are up to 2 people in the room, the camera also has the ability to augment the UWB measurements by roughly 15%. As the number of people in the room varies from 3-6, the precision of the LADAR becomes worse and the error increases from 20 to 30 cm. However, the precision is still better than the UWB. The LADAR can therefore still augment the precision of the UWB. When the number of people is greater than 6, the precision becomes worse than the UWB and the LADAR is no longer able to improve the precision of the fusion. The camera's precision is worse than that of the UWB when the number of people is greater than 2. In conclusion, when the number of people in the room is greater than 6 these particular unaided sensors did not result in improvement in tracking precision over the UWB sensor.

Chapter 4

Conclusions and Future Work

UWB indoor position tracking has the potential to enable new applications in telepresence, virtual reality, training, asset tracking, navigation and entertainment. However, the current room/building level precision is on the order of 30 cm. This needs to be improved to near 1 cm (or better) to enable these applications. Inspired by methods that have been used to improve GPS tracking, this dissertation has explored novel techniques to improve UWB tracking precision through noise modeling and augmentation.

In chapter 2 we identified a new noise source due to the switching of fixed point sets for trilateration. While this noise is theoretically present in all trilateration-based systems, it is not readily apparent in large-scale systems like the GPS, but it can cause noticeable jump-like behavior in indoor UWB position tracking. We developed a mathematical model and particle filter that accounts for this noise and tested our methods on a real UWB indoor position tracking system. Our set noise particle filter showed an approximately 15% improvement in accuracy over the raw measurements. While this improvement is useful, it is the opinion of this author that noise due to NLOS and multipath errors tends to be more significant than noise caused by sensor

set switching. Changes in the orientation of the transmitter tags also cause a varying noise. Future work should focus on exploring the effects of these noise sources, and possibly others. The methodology proposed in this dissertation could be used to approach each of these noises in isolation or in combination.

In chapter 3 we explored the idea of augmenting UWB with a differential sensor. Two types of differential sensor were tested, a single LADAR and a network of 6 cameras operating to created a floor-level occupancy map. Using LADAR augmentation improved MOTA for 1-8 people in a room, and improved MOTP for 1-6 people by as much as 40% at 1 person. Using camera augmentation improved MOTA for 1-8 people in a room and improved MOTP for 1 person by 15%, but decreased MOTP for 3+ people. Future work should explore sensor noise models that vary depending on the number of people in the room. A more advanced noise model could incorporate information about the occlusions of the differential sensor caused by people as they move. Other types of differential sensors could be tried, such as a single LOS UWB receiver. Another augmentation that could be explored is the inclusion of a gyroscope or IMU to track the orientation of the transmitting tag, in order to model and mitigate the effect of antenna orientation noise.

In the process of exploring differential augmentation for UWB position tracking, we developed the idea of fusing data from an aided and unaided sensor. This framework could be applied to other problems. For example, body tracking using camera and body-worn MEMS sensors might be enhanced by our data fusion filter framework.

The ultimate goal of this research is to push towards centimeter level accuracy in building-sized indoor tracking. It took over 30 years for GNSS to progress from battlespace tracking to consumer level applications, and one would expect that indoor position tracking will take a similar amount of time to progress from military research

systems to consumer applications. The work described in this dissertation should be viewed in that context as only one step in that direction.

Appendices

Appendix A

Curriculum Vitae

Salil Partha Banerjee

105 College Street #22
Clemson, SC 29631
864-508-4042
salilb@g.clemson.edu

EDUCATION

Ph.D. in Electrical Engineering, December 2012
Clemson University, Clemson

B.S. in Electronics Engineering, May 2008
D. J. S. College of Engineering, University of Mumbai, Mumbai

PUBLICATIONS

- S. Banerjee, W. Suski, A. Hoover, *"Identifying and Filtering Noise Caused by Sensor Set Switching in UWB Indoor Position Tracking"*, International Journal of Ultra Wideband Communications and Systems (Under review)
- W. Suski, S. Banerjee, A. Hoover, *"Using a Map of Measurement Noise Map to Improve UWB Indoor Position Tracking"*, IEEE Transactions on Instrumentation and Measurement (accepted)
- Y. Dong, S. Banerjee, D. Woodard, *"Augmenting Iris with Eyebrow Features to Improve Biometric Recognition and Gender Classification for Non-Ideal Images"* (In progress for submission)

- S. Banerjee, D. Woodard, *"Biometric Authentication and Identification using Keystroke Dynamics: A Survey"*, Journal of Pattern Recognition Research, vol. 7, no.1, pp. 116-139, 2012
- S. Banerjee, W. Suski, A. Hoover, *"Sensor Set Switching in UWB Indoor Position Tracking"*, International Conference on Ultra-Wideband, Sept. 17-20, 2012
- W. Suski, S. Banerjee, A. Hoover, *"System-Level Noise of An Ultra-Wideband Tracking System"*, The 11th International Conference on Information Science, Signal Processing and their Applications, pp. 634-639, July 2-5, 2012

PROFESSIONAL EXPERIENCE

Graduate Research Assistant
ECE Department, Clemson University, Clemson, SC
UWB-based Indoor Positioning System / MOUT (dissertation work)
Aug 2012 - Nov 2012
- Identified and model noise sources in ultra-wideband positioning systems
- Augmented UWB with Ladar and cameras to improve accuracy
- Improved tracking by using Kalman and particle filtering techniques
- Funded in part by Office of Naval Research

Graduate Research Assistant
BPRL, School of Computing at Clemson University, Clemson, SC
Jan 2012 - Nov 2012
- Developed algorithms to automatically detect eye corners in facial images
- Developed models to detect key features around periocular regions for biometric
- Funded by Federal Bureau of Investigation

Graduate Research Assistant
ECE Department, Clemson University, Clemson, SC
Head Mounted Display (HMD) Lag Project
Aug 2011 - Dec 2011
- Designed a test bench to verify delay model in a HMD
- Developed software to vary lag and measure effect on humans

Graduate Research Assistant
BPRL, School of Computing at Clemson University, Clemson, SC
Keystroke Biometrics
Apr 2011 - Aug 2011
- Conducted preliminary study to identify gender using typing cadence
- Developed models to identify users based on keystroke dynamics

Graduate Teaching Assistant
ECE Department, Clemson University, Clemson, SC
Jan 2011 - Apr 2011
- Laboratory Instructor for the undergraduate course Embedded Computing
- Set up lab, supervise and grade assignments

Graduate Grading Assistant
ECE Department, Clemson University, Clemson, SC
Jan 2009 - Apr 2009
- Served as a grader for the course System Programming Concepts
- Graded homework and test papers

Graduate Teaching Assistant
ECE Department, Clemson University, Clemson, SC
Aug 2008 - Dec 2008
- Laboratory Instructor for the undergraduate course Computer Organization
- Instructed, supervised and graded papers for a group of students

Undergraduate Research Assistant
NMIMS University, Mumbai, India
Detection of Emotional Expression from Speech
Aug 2007 - Apr 2008
- Created a speech database with different emotions
- Detected emotional expressions in speech autonomously from data
- Compared voice parameters extracted from speech for emotion recognition

TECHNICAL SKILLS

- **Languages:** C, Win 32 API, C++, Core Java, MATLAB, SIMULINK
- **Packages:** OpenCV, Blepo
- **Statistical software:** SAS, R (GNU S)
- **Tools:** LaTex, CodeVision AVR, Pinnacle
- **Web Technologies:** HTML
- **Platforms:** Linux, Windows, Mac

ACADEMIC HONORS

- Earned academic scholarships from J.R.D. Tata Trust for excellence in Mumbai University, India, 2001-2004
- Ranked 8th in Maharashtra State (Merit List - Mumbai Division), India in High School, 2000

LEADERSHIP EXPERIENCE

- President of Association for India's Development (AID), Clemson chapter, 2011-2012
- Senator representing Electrical Engineering Department at Clemson University, 2009-2011
- Led the Web Designing Competition at IEEE technical festival, 2007
- Led the Robotics event at IETE technical festival, 2007

MEMBERSHIPS

- Omicron Delta Kappa (ODK) Honor Society
- Institute of Electrical and Electronics Engineers (IEEE)
- Institute of Navigation (ION)

Bibliography

[1] F. Ahmad, Y. Zhang, and M. Amin. Three-dimensional wideband beamforming for imaging through a single wall. *IEEE Geoscience and Remote Sensing Letters*, 5(2):176–179, 2008.

[2] N. Alsindi and K. Pahlavan. Cooperative localization bounds for indoor ultra-wideband wireless sensor networks. *EURASIP Journal on Advances in Signal Processing*, 2008(1), 2007.

[3] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2):174–188, February 2002.

[4] Y. Bar-Shalom and X. Li. *Estimation and Tracking - Principles, Techniques, and Software*. Artech House, 1993.

[5] Y. Bar-Shalom, X. Li, and T. Kirubarajan. *Estimation with Applications to Tracking and Navigation: Theory Algorithms and Software*. Wiley-Interscience, 2001.

[6] Y. Bar-Shalom and T.E.Fortmann. *Tracking and Data Association*. Academic Press, 1988.

[7] G. Bebis, A. Gyaourova, S. Singh, and I. Pavlidis. Face recognition by fusing thermal infrared and visible imagery. *Image and Vision Computing*, 24(7):727–742, 2006.

[8] K. Bernardin, A. Elbs, and R. Stiefelhagen. Multiple object tracking performance metrics and evaluation in a smart room environment. In *Sixth IEEE International Workshop on Visual Surveillance, in conjunction with ECCV*, 2006.

[9] K. Bernardin and R. Stiefelhagen. Evaluating multiple object tracking performance: the CLEAR MOT metrics. *Journal on Image and Video Processing*, 2008:1, 2008.

[10] D. Beymer. Person counting using stereo. In *Workshop on Human Motion*, pages 127–133. IEEE, 2000.

[11] C. Bilich. Bio-medical sensing using ultra wideband communications and radar technology: A feasibility study. In *Pervasive Health Conference and Workshops*, pages 1–9. IEEE, 2006.

[12] J. Black, T. Ellis, and P. Rosin. A novel method for video tracking performance evaluation. In *International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance*, pages 125–132, 2003.

[13] T. Bosch and M. Lescure. Experimental determination of the useful reflection coefficient of non-cooperative targets for a time-of-flight laser rangefinderx. *Optical Review*, 2:289–291, 1995.

[14] R. Brown. *Introduction to random signal analysis and Kalman filtering*, volume 8. John Wiley & Sons Inc, 1983.

[15] B. Brumitt, B. Meyers, J. Krumm, A. Kern, and S. Shafer. Easyliving: Technologies for intelligent environments. In P. Thomas and H.-W. Gellersen, editors, *Handheld and Ubiquitous Computing*, volume 1927 of *Lecture Notes in Computer Science*, pages 97–119. 2000.

[16] M. Campbell and W. Whitacre. Cooperative tracking using vision measurements on SeaScan UAVs. *IEEE Transactions on Control Systems Technology*, 15(4):613–626, July 2007.

[17] F. Caron, M. Davy, E. Duflos, and P. Vanheege. Particle filtering for multisensor data fusion with switching observation models: Application to land vehicle positioning. *IEEE Transactions on Signal Processing*, 55(6):2703–2719, June 2007.

[18] H. Carvalho, W. Heinzelman, A. Murphy, and C. Coelho. A general data fusion architecture. In *International Confernce on Information Fusion*, volume 2, pages 1465–1472, 2003.

[19] R. Chandra, A. Gaikwad, D. Singh, and M. Nigam. An approach to remove the clutter and detect the target for ultra-wideband through-wall imaging. *Journal of Geophysics and Engineering*, 5(4):412, 2008.

[20] X. Chen, J. Liang, S. Wang, Z. Wang, and C. Parini. Small ultra wideband antennas for medical imaging. In *Antennas and Propagation Conference*, pages 28–31. IEEE, 2008.

[21] K. Cheok, M. Radovnikovich, P. Vempaty, G. Hudas, J. Overholt, and P. Fleck. UWB tracking of mobile robots. In *IEEE 21st International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, pages 2615–2620, September 2010.

[22] CLEAR. Classification of events, activities and relationships (CLEAR) workshop. http://clear-evaluation.org/.

[23] CNYPhotoboothRental. Motion capture suit. http://cnyphotobooth.com/3d-stereogram-photo-booth.htm.

[24] R. Collins, A. Lipton, H. Fujiyoshi, and T. Kanade. Algorithms for cooperative multisensor surveillance. *Proceedings of the IEEE*, 89(10):1456–1477, October 2001.

[25] J. A. Corrales, F. A. Candelas, and F. Torres. Hybrid tracking of human operators using IMU/UWB data fusion by a Kalman filter. In *Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction*, HRI '08, pages 193–200, 2008.

[26] B. Dasarathy. Sensor fusion potential exploitation-innovative architectures and illustrative applications. *Proceedings of the IEEE*, 85(1):24–38, January 1997.

[27] S. Davis, B. Van Veen, S. Hagness, and F. Kelcz. Breast tumor characterization based on ultrawideband microwave backscatter. *IEEE Transactions on Biomedical Engineering*, 55(1):237–246, 2008.

[28] B. Denis, L. Ouvry, B. Uguen, and F. Tchoffo-Talom. Advanced Bayesian filtering techniques for UWB tracking systems in indoor environments. In *IEEE International Conference on Ultra-Wideband*, page 6, Septemnber 2005.

[29] P. Djuric, J. Kotecha, J. Zhang, Y. Huang, T. Ghirmai, M. Bugallo, and J. Miguez. Particle filtering. *IEEE Signal Processing Magazine*, 20(5):19–38, September 2003.

[30] A. S. Dmitriev, B. E. Kyarginskii, A. I. Panas, D. Y. Puzikov, and S. O. Starkov. Ultrawideband direct chaotic data transmission in the microwave range. *Technical Physics Letters*, 29(1):72–74, 2003.

[31] J. Dong, D. Zhuang, Y. Huang, and J. Fu. Advances in multi-sensor data fusion: algorithms and applications. *Sensors*, 9(10):7771–7784, September 2009.

[32] H. Durrant-Whyte. Multi sensor data fusion. Technical report, Australian Center for Field Robotics, The University of Sydney, NSW, 2001.

[33] Ekahau. Comparison of wireless indoor positioning technologies. http://www.productivet.com/docs-2/Wireless_Comparison.pdf.

[34] A. Ekimov and J. M. Sabatier. Human detection range by active doppler and passive ultrasonic methods. In E. M. Carapezza, editor, *Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense VII*, volume 6943. SPIE, 2008.

[35] A. Elkins, E. Muth, A. Hoover, A. Walker, T. Carpenter, and F. Switzer. Physiological compliance and team performance. *Applied Ergonomics*, 40:997–1003, 2009.

[36] T. Ellis. Performance metrics and methods for tracking in surveillance. In *IEEE International Workshop on Performance Evaluation of Tracking and Surveillance*, 2002.

[37] P. Enge, R. Kalafus, and M. Ruane. Differential operation of the global positioning system. *IEEE Communications Magazine*, 26(7):48–60, July 1988.

[38] A. Farina, P. Lombardo, and M. Marsella. Joint tracking and identification algorithms for multisensor data. *IEE Proceedings - Radar, Sonar and Navigation*, 149(6):271–280, December 2002.

[39] J. Farrell. *Aided navigation: GPS with high rate sensors.* McGraw-Hill Companies, 2008.

[40] FCC 02-48. In the matter of revision of part 15 of the commission's rule regarding ultra-wideband transmission systems. Technical report, Federal Communications Commission, April 2002.

[41] R. Fontana. Recent system applications of short-pulse ultra-wideband (UWB) technology. *IEEE Transactions on Microwave Theory and Techniques*, 52(9), September 2004.

[42] R. Fontana and S. Gunderson. Ultra-wideband precision asset location system. In *IEEE Conference on Ultra Wideband Systems and Technologies*, pages 147–150. IEEE, 2002.

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

[44] C. Fowler, J. Entzminger, and J. Corum. Assessment of ultra-wideband (UWB) technology, year=1990, month=November, volume=5, number=11, pages=45–49. *IEEE Aerospace and Electronic Systems Magazine.*

[45] D. Ganjali. Filtering noise caused by sensor selection for an ultra-wideband position tracking system. Master's thesis, Clemson University, December 2009.

[46] J. Gavan. Transponders for the detection and identification of remote cooperative targets. In *IEEE National Telesystems Conference*, pages 229–232, May 1994.

[47] C. Gentile and A. Kik. A comprehensive evaluation of indoor ranging using ultra-wideband technology. *EURASIP Journal on Wireless Communication Networks*, 2007:1–10, January 2007.

[48] J. Georgy, T. Karamat, U. Iqbal, and A. Noureldin. Enhanced MEMS-IMU/odometer/GPS integration using mixture particle filter. *GPS Solutions*, 15:239–252, 2011.

[49] S. Gezici, Z. Tian, G. Giannakis, H. Kobayashi, A. Molisch, H. Poor, and Z. Sahinoglu. Localization via ultra-wideband radios. *IEEE Signal Processing Magazine*, 22(4):70–84, 2005.

[50] M. Ghavami, L. Michael, and R. Kohno. *Ultra Wideband Signals and Systems in Communication Engineering*. Wiley, 2 edition, 2007.

[51] A. Ghildiyal, K. Amara, R. Molin, B. Godara, A. Amara, and R. Shevgaonkar. UWB for in-body medical implants: a viable option. In *IEEE International Conference on Ultra-Wideband (ICUWB)*, volume 2, pages 1–4, September 2010.

[52] M. Grewal, L. Weill, and A. Andrews. *Global Positioning Systems, Inertial Navigation, and Integration*. Wiley-Interscience, 2 edition, 2007.

[53] Z. Guoping and S. Rao. Position localization with impulse ultra wide band. In *IEEE/ACES International Conference on Wireless Communications and Applied Computational Electromagnetics*, pages 17–22, April 2005.

[54] F. Gustafsson, F. Gunnarsson, N. Bergman, U. Forssell, J. Jansson, R. Karlsson, and P.-J. Nordlund. Particle filters for positioning, navigation, and tracking. *IEEE Transactions on Signal Processing*, 50(2):425–437, February 2002.

[55] D. Hall. *Mathematical Techniques in Multisensor Data Fusion*. Artech House, 1st edition edition, 1992.

[56] D. Hall and J. Llinas. An introduction to multisensor data fusion. *Proceedings of the IEEE*, 85(1):6–23, January 1997.

[57] A. Harter, A. Hopper, P. Steggles, A. Ward, and P. Webster. The anatomy of a context-aware application. *Wireless Networks*, 8:187–197, 2002.

[58] M. Hernandez and R. Kohno. UWB systems for body area networks in IEEE 802.15.6. In *IEEE International Conference on Ultra-Wideband (ICUWB)*, pages 235–239, September 2011.

[59] M. Hernandez-Pajares, J. Zomoza, J. Subirana, and O. Colombo. Feasibility of wide-area subdecimeter navigation with GALILEO and Modernized GPS. *IEEE Transactions on Geoscience and Remote Sensing*, 41(9):2128–2131, September 2003.

[60] Hexamite. Hx17 network. http://www.hexamite.com/hx17a.htm.

[61] T. W. Highlands and S. C. A. Thomopoulos. Detection and range/doppler estimation for colocated sensors. *IEEE Transactions on Aerospace and Electronic Systems*, 33(3):825–834, July 1997.

[62] B. Hofmann-Wellenhof, H. Lichtenegger, and J. Collins. *Global Positioning System: Theory and Practice*. Springer, 4 edition, 1997.

[63] V. Honkavirta, T. Perala, S. Ali-Loytty, and R. Piche. A comparative survey of WLAN location fingerprinting methods. In *6th Workshop on Positioning, Navigation and Communication*, pages 243–251, March 2009.

[64] A. Hoover. Lecture notes - ECE 854, 2008.

[65] A. Hoover and E. Muth. *The PSI Handbook of Virtual Environments for Training and Education, Volume 3:Integrated Systems, Training Evaluations and Future Directions*, chapter Instrumenting for measuring, pages 184–195. Praeger Security International Publishing, 2008.

[66] A. Hoover and B. Olsen. A real-time occupancy map from multiple video streams. In *IEEE International Conference on Robotics and Automation*, volume 3, pages 2261–2266. IEEE, 1999.

[67] P. Hyde, R. Dubayah, W. Walker, J. Blair, M. Hofton, and C. Hunsaker. Mapping forest structure for wildlife habitat analysis using multi-sensor (LiDAR, SAR/InSAR, ETM+, Quickbird) synergy. *Remote Sensing of Environment*, 102(1):63–73, 2006.

[68] M. Jing, Z. Nai-tong, and Z. Qin-yu. IR-UWB waveform distortion analysis in NLOS localization system. *Information Technology Journal*, 9:139–145, 2010.

[69] D. Jourdan, J. J. Deyst, M. Win, and N. Roy. Monte Carlo localization in dense multipath environments using UWB ranging. In *IEEE International Conference on Ultra-Wideband*, pages 314–319, September 2005.

[70] K. Kaemarungsi and P. Krishnamurthy. Modeling of indoor positioning systems based on location fingerprinting. In *Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies*, volume 2, pages 1012–1022, March 2004.

[71] E. Kaplan and C. J. Hegarty, editors. *Understanding GPS, principles and applications*. Artech House, Inc., 2006.

[72] R. Kasturi, D. Goldgof, P. Soundararajan, V. Manohar, M. Boonstra, and V. Korzhova. Performance evaluation protocol for face, person and vehicle detection & tracking in video analysis and content extraction (VACE-II). Technical report, Submitted to Advanced Research and Development Activity, 2006.

[73] C. Kee and B. Parkinson. Wide area differential GPS (WADGPS): future navigation system. *IEEE Transactions on Aerospace and Electronic Systems*, 32(2):795–808, April 1996.

[74] A. Kelly. A 3D state space formulation of a navigation Kalman filter for autonomous vehicles. Technical report, DTIC Document, 1994.

[75] J.-A. Kim, E. W. Bae, S. H. Kim, and Y. K. Kwak. Design methods for six-degree-of-freedom displacement measurement systems using cooperative targets. *Precision Engineering*, 26(1):99–104, 2002.

[76] I. D. Kitching. GPS and cellular radio measurement integration. *The Journal of Navigation*, 53(03):451–463, 2000.

[77] R. Knoblauch, M. Pietrucha, and M. Nitzburg. Field studies of pedestrian walking speed and start-up time. *Transportation Research Record*, 1538(1):27–38, 1996.

[78] K. W. Kolodziej and J. Hjelm. *Local positioning systems: LBS applications and services*. CRC Press, 2006.

[79] S. Kong, J. Heo, B. Abidi, J. Paik, and M. Abidi. Recent advances in visual and infrared face recognition: A review. *Computer Vision and Image Understanding*, 97(1):103–135, 2005.

[80] J. Kotecha and P. Djuric. Gaussian particle filtering. *IEEE Transactions on Signal Processing*, 51(10):2592–2601, October 2003.

[81] M. Kuhn, C. Zhang, B. Merkl, D. Yang, Y. Wang, M. Mahfouz, and A. Fathy. High accuracy UWB localization in dense indoor environments. In *IEEE International Conference on Ultra-Wideband (ICUWB)*, volume 2, pages 129–132, September 2008.

[82] I. Labs. Following the idf: Ultra wide band wireless data transfer technology. http://ixbtlabs.com/articles2/uwb/index.html.

[83] C. Lai and R. Narayanan. Through-wall imaging and characterization of human activity using ultrawideband (UWB) random noise radar. In *Proceedings of SPIE*, volume 5778, pages 186–195, 2005.

[84] M. Lazebnik, E. Madsen, G. Frank, and S. Hagness. Tissue-mimicking phantom materials for narrowband and ultrawideband microwave applications. *Physics in Medicine and Biology*, 50(18):4245, 2005.

[85] W. Lechner and S. Baumann. Global navigation satellite systems. *Computers and Electronics in Agriculture*, 25(1-2):67–85, 2000.

[86] J.-C. Lementec and P. Bajcsy. Recognition of arm gestures using multiple orientation sensors: gesture classification. In *The 7th International IEEE Conference on Intelligent Transportation Systems*, pages 965–970, October 2004.

[87] L. Letham. *GPS made easy: using global positioning systems in the outdoors*, page 224. Mountaineers Books, 3 edition, November 2003.

[88] F. Lewis. *Optimal Estimation with an Introduction to Stochastic Control Theory*. Wiley-Interscience, 1986.

[89] X. Li and V. Jilkov. A survey of maneuvering target tracking: Dynamic models. In *Proceedings of SPIE Conference on Signal and Data Processing of Small Targets*, volume 4048, pages 212–236, 2000.

[90] J. Lie, C. M. See, and B. P. Ng. Uwb ranging with high robustness against dominant jammer and multipath. *Microwave and Wireless Components Letters, IEEE*, 15(12):907–909, December 2005.

[91] J. H. Lim, I.-J. Wang, and A. Terzis. Tracking a non-cooperative mobile target using low-power pulsed doppler radars. In *IEEE 35th Conference on Local Computer Networks (LCN)*, pages 913–920, October 2010.

[92] H. Liu, M. Bolic, A. Nayak, and I. Stojmenovic. Taxonomy and challenges of the integration of RFID and wireless sensor networks. *IEEE Network*, 22(6):26–35, November 2008.

[93] H. Liu, H. Darabi, P. Banerjee, and J. Liu. Survey of wireless indoor positioning techniques and systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 37(6):1067–1080, November 2007.

[94] W. Liu, H. Ding, X. Huang, X. Li, and J. Yuan. Preliminary study on noncooperative positioning using UWB impulse radio. In *IEEE International Conference on Ultra-wideband*, September 2012.

[95] Y. Liu, A. Hoover, and I. Walker. A timing model for vision-based control of industrial robot manipulators. *IEEE Transactions on Robotics*, 20(5):891–898, October 2004.

[96] Z. Low, J. Cheong, C. Law, W. Ng, and Y. Lee. Pulse detection algorithm for line-of-sight (LOS) UWB ranging applications. *IEEE Antennas and Wireless Propagation Letters*, 4:63–67, 2005.

[97] R. Luo, C. Yih, and K. Su. Multisensor fusion and integration: approaches, applications, and future research directions. *Sensors Journal, IEEE*, 2(2):107–119, April 2002.

[98] R. C. Luo and M. G. Kay. *Data Fusion in Robotics and Machine Intelligence*, chapter Data fusion and sensor integration: state-of-the-art 1990s, pages 7–136. Academic Press, 1992.

[99] G. MacGougan, K. O'Keefe, and R. Klukas. Tightly-coupled GPS/UWB integration. *The Journal of Navigation*, 63(01):1–22, 2010.

[100] M. Mahfouz, C. Zhang, B. Merkl, M. Kuhn, and A. Fathy. Investigation of high-accuracy indoor 3-D positioning using UWB technology. *IEEE Transactions on Microwave Theory and Techniques*, 56(6):1316 –1330, June 2008.

[101] V. Manohar, P. Soundararajan, H. Raju, D. Goldgof, R. Kasturi, and J. Garofolo. Performance evaluation of object detection and tracking in video. *Computer Vision–ACCV 2006*, pages 151–161, 2006.

[102] R. Mautz. Overview of current indoor positioning systems. In *Geodesy and Cartography*, volume 34, pages 66–70, 2009.

[103] K. Mechitov, S. Sundresh, Y. Kwon, and G. Agha. Cooperative tracking with binary-detection sensor networks. In *SenSys*, pages 332–333, November 2003.

[104] C. Meier, A. Terzis, and S. Lindenmeier. A robust 3D high precision radio location system. In *IEEE/MTT-S International Microwave Symposium*, pages 397–400, June 2007.

[105] L. E. Miller. Why UWB? A review of ultrawideband technology. Technical report, Report to NETEX Project Office, DARPA, 2003.

[106] P. Misra and P. Enge. *Global Positioning System: Signals, Measurements and Performance*. Ganga-Jamuna Press, 2001.

[107] R. Möbus and U. Kolbe. Multi-target multi-object tracking, sensor fusion of radar and infrared. In *IEEE Intelligent Vehicles Symposium*, pages 732–737. IEEE, 2004.

[108] T. B. Moeslund and E. Granum. A survey of computer vision-based human motion capture. *Computer Vision and Image Understanding*, 81(3):231–268, 2001.

[109] A. Muqaibel, A. Safaai-Jazi, A. Bayram, A. Attiya, and S. Riad. Ultrawideband through-the-wall propagation. *IEE Proceedings - Microwaves, Antennas & Propagation*, pages 581–588, December 2005.

[110] S. Nag, M. Barnes, T. Payment, and G. Holladay. Ultrawideband through-wall radar for detecting the motion of people in real time. In *AeroSense 2002*, pages 48–57. International Society for Optics and Photonics, 2002.

[111] S. Nandulal, C. Babu Rao, C. Indi, M. Irulappan, S. Arulmozhi, and P. Soma. Evaluation of real-time position accuracy and LNAV/VNAV service availability of GAGAN SBAS (Wide Area Differential GPS) over Indian region. In *Tyrrhenian International Workshop on Digital Communications - Enhanced Surveillance of Aircraft and Vehicles*, pages 1–6, September 2008.

[112] M. Navarro and M. Najar. TOA and DOA estimation for positioning and tracking in IR-UWB. In *IEEE International Conference on Ultra-Wideband (ICUWB)*, pages 574–579, September 2007.

[113] A. Nghiem, F. Bremond, M. Thonnat, and V. Valentin. Etiseo, performance evaluation for video surveillance systems. In *IEEE Conference on Advanced Video and Signal Based Surveillance*, pages 476–481. IEEE, 2007. http://www-sop.inria.fr/orion/ETISEO/.

[114] L. M. Ni, Y. Liu, Y. C. Lau, and A. P. Patil. LANDMARC: indoor location sensing using active RFID. *Wireless Networks*, 10:701–710, 2004.

[115] I. Oppermann, M. Hmlinen, and J. Iinatti, editors. *UWB: Theory and applications*. Wiley, November 2004.

[116] R. J. Orr and G. D. Abowd. The smart floor: a mechanism for natural user identification and tracking. In *Extended Abstracts on Human factors in Computing Systems*, CHI EA '00, pages 275–276, 2000.

[117] B. Parkinson and J. Spilker, editors. *Global Positioning System: Theory and Applications*, volume 1. American Institute of Aeronautics and Astronautics (AIAA), January 1996.

[118] PETS. PETS: Performance evaluation of tracking and surveillance workshop. http://www.cvg.rdg.ac.uk/slides/pets.html, 2007.

[119] S. Pittet, V. Renaudin, B. Merminod, and M. Kasser. UWB and MEMS based indoor navigation. *The Journal of Navigation*, 61(03):369–384, 2008.

[120] C. Pohl and J. Van Genderen. Review article multisensor image fusion in remote sensing: concepts, methods and applications. *International journal of remote sensing*, 19(5):823–854, 1998.

[121] D. Porcino and W. Hirt. Ultra-wideband radio technology: potential and challenges ahead. *IEEE Communications Magazine*, 41(7):66–74, 2003.

[122] N. B. Priyantha, A. Chakraborty, and H. Balakrishnan. The Cricket location-support system. In *Proceedings of the 6th annual international conference on Mobile computing and networking*, MobiCom '00, pages 32–43, 2000.

[123] H. Qi and J. B. Moore. Direct Kalman filtering approach for GPS/INS integration. *IEEE Transactions on Aerospace and Electronic Systems*, 38(2):687 –693, April 2002.

[124] B. R. R. and I. S. *Multi-Sensor Fusion: Fundamentals and Applications with Software*. Prentice Hall PTR, 1st edition edition, 1997.

[125] Rafael Muñoz-Salinas and E. Yeguas-Bolivar and L. Díaz-Más and Rafael Medina-Carnicer. Shape from pairwise silhouettes for plan-view map generation. *Image and Vision Computing*, 30(2):122–133, 2012.

[126] I. M. Rekleitis. A particle filter tutorial for mobile robot localization, Technical Report TR-CIM-04-02. Technical report, Centre for Intelligent Machines, McGill University, 2004.

[127] M. Rudoy, C. Rohrs, and J. Chen. Signatures of walking humans from passive and active acoustic data using time-varying vector autoregressions. In *Conference Record of the Forty-First Asilomar Conference on Signals, Systems and Computers*, pages 2253–2256, November 2007.

[128] Z. Sahinoglu, S. Gezici, and I. Guvenc. *Ultra-Wideband Positioning Systems*. Cambridge University Press, 2 edition, October 2008.

[129] U. Scheunert, H. Cramer, B. Fardi, and G. Wanielik. Multi sensor based tracking of pedestrians: a survey of suitable movement models. In *IEEE Intelligent Vehicles Symposium*, pages 774–778, June 2004.

[130] G. Shen, R. Zetik, O. Hirsch, and R. Thomä. Range-based localization for UWB sensor networks in realistic environments. *EURASIP Journal on Wireless Communications and Networking*, 2010.

[131] SICK. LMS 200/LMS 211/LMS 220/LMS 221/LMS 291 Laser measurement systems. http://www.sick-automation.ru/images/File/pdf/LMS%20Technical%20Description.pdf.

[132] F. Siegemund and C. Florkemeier. Interaction in pervasive computing settings using Bluetooth-enabled active tags and passive RFID technology together with mobile phones. In *Proceedings of the First IEEE International Conference on Pervasive Computing and Communications*, pages 378–387, March 2003.

[133] M. Soliman, T. Morimoto, and Z. Kawasaki. Three-dimensional localization system for impulsive noise sources using ultra-wideband digital interferometer technique. *Journal of Electromagnetic Waves and Applications*, 20(4):515–530, 2006.

[134] S.S.Blackman. *Multitarget-Multisensor Tracking: Advanced Applications*, chapter Association and Fusion of Multiple Sensor Data. Artech House, 1990.

[135] R. Stiefelhagen, K. Bernardin, R. Bowers, J. Garofolo, D. Mostefa, and P. Soundararajan. The CLEAR 2006 evaluation. In R. Stiefelhagen and J. Garofolo, editors, *Multimodal Technologies for Perception of Humans*, volume 4122 of *Lecture Notes in Computer Science*, pages 1–44. Springer Berlin / Heidelberg, 2007.

[136] R. Stiefelhagen, K. Bernardin, R. Bowers, R. Rose, M. Michel, and J. Garofolo. The CLEAR 2007 evaluation. In R. Stiefelhagen, R. Bowers, and J. Fiscus, editors, *Multimodal Technologies for Perception of Humans*, volume 4625 of *Lecture Notes in Computer Science*, pages 3–34. Springer Berlin / Heidelberg, 2008.

[137] R. Stiefelhagen, H. Steusloff, and A. Waibel. CHIL-computers in the human interaction loop. In *Proceedings of the International Workshop on Image Analysis for Multimedia Interactive Services*, 2004.

[138] G. Sun, J. Chen, W. Guo, and K. Liu. Signal processing techniques in network-aided positioning: a survey of state-of-the-art positioning designs. *IEEE Signal Processing Magazine*, 22(4):12–23, July 2005.

[139] W. Suski. *A Study of Environment Noise in Ultra-Wideband Indoor Position Tracking*. PhD thesis, Clemson University, May 2012.

[140] W. Suski, S. Banerjee, and A. Hoover. System-level noise of an ultra-wideband tracking system. In *IEEE The* 11^{th} *International Conference on Information Science, Signal Processing and their Applications*, July 2012.

[141] N. Swangmuang and P. Krishnamurthy. Location fingerprint analysis toward efficient indoor positioning. In *Sixth Annual IEEE International Conference on Pervasive Computing and Communications*, pages 100–109, March 2008.

[142] A. Thiagarajan, J. Biagioni, T. Gerlich, and J. Eriksson. Cooperative transit tracking using smart-phones. In *Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems*, SenSys '10, pages 85–98, November 2010.

[143] V. Tsagaris and V. Anastassopoulos. Fusion of visible and infrared imagery for night color vision. *Displays*, 26(4):191–196, 2005.

[144] Ubisense. Series 7000 compact tag. http://www.ubisense.net/en/resources/factsheets/series-7000-compact-tag.html.

[145] Ubisense. Series 7000 sensor. http://www.ubisense.net/en/resources/factsheets/series-7000-sensor.html.

[146] Ubisense. Ubisense precise location. http://www.ubisense.net/en/resources/factsheets/ubisense-precise-location.html.

[147] R. Van der Merwe and E. Wan. The square-root unscented Kalman filter for state and parameter-estimation. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 6, pages 3461–3464, 2001.

[148] P. K. Varshney. Multisensor data fusion. *Electronics & Communication Engineering Journal*, 9(6):245–253, December 1997.

[149] A. Waibel and R. Stiefelhagen. *Computers in the Human Interaction Loop*. Springer, 2 edition, 2009.

[150] L. Wald. Some terms of reference in data fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 37(3):1190–1193, May 1999.

[151] E. Wan and R. Van Der Merwe. The unscented Kalman filter for nonlinear estimation. In *The IEEE Adaptive Systems for Signal Processing, Communications, and Control Symposium (AS-SPCC)*, pages 153–158, 2000.

[152] R. L. Waters and S. Mulroy. The energy expenditure of normal and pathologic gait. *Gait & Posture*, 9(3):207–231, 1999.

[153] G. Welch and G. Bishop. An introduction to the kalman filter, 1995.

[154] G. Welch and G. Bishop. SCAAT: Incremental Tracking with Incomplete Information. In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*, pages 333–344. ACM PressAddison-Wesley Publishing Co., 1997.

[155] F. White. Data fusion lexicon. Technical report, DTIC Document, 1991.

[156] H. Wymeersch, J. Lien, and M. Win. Cooperative localization in wireless networks. *Proceedings of the IEEE*, 97(2):427–450, February 2009.

[157] N. Xiong and P. Svensson. Multi-sensor management for information fusion: issues and approaches. *Information fusion*, 3(2):163–186, September 2002.

[158] D. Yang, A. Fathy, H. Li, M. Mahfouz, and G. Peterson. Millimeter accuracy UWB positioning system using sequential sub-sampler and time difference estimation algorithm. In *IEEE Radio and Wireless Symposium (RWS)*, pages 539–542, January 2010.

[159] D. Young, C. Keller, D. Bliss, and K. Forsythe. Ultra-wideband (UWB) transmitter location using time difference of arrival (TDOA) techniques. In *Signals, Systems and Computers, 2003. Conference Record of the Thirty-Seventh Asilomar Conference on*, volume 2, pages 1225–1229. IEEE, 2003.

[160] R. Zetik, J. Sachs, and R. Thoma. UWB localization - active and passive approach [ultra wideband radar]. In *Proceedings of the 21st IEEE Instrumentation and Measurement Technology Conference*, volume 2, pages 1005–1009, May 2004.

[161] A. Zhalilo and V. Shokalo. Status and prospects of differential navigation and high precision positioning GNSS-technologies in Ukraine. In *International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science (TCSET)*, pages 6–9, February 2010.

[162] C. Zhang, M. Kuhn, B. Merkl, A. Fathy, and M. Mahfouz. Preal-time noncoherent uwb positioning radar with millimeter range accuracy: theory and experiment. *IEEE Transactions on Microwave Theory and Techniques*.

[163] C. Zhang, M. Kuhn, B. Merkl, M. Mahfouz, and A. Fathy. Development of an UWB indoor 3D positioning radar with millimeter accuracy. In *IEEE MTT-S International Microwave Symposium Digest*, pages 106–109, June 2006.

[164] F. Zhang, J. Chen, H. Li, Y. Sun, and X. Shen. Distributed interfering sensor scheduling scheme for target tracking. In *5th International ICST Conference on Communications and Networking in China (CHINACOM)*, pages 1–6, August 2010.

[165] Y. Zhang and Q. Ji. Active and dynamic information fusion for multisensor systems with dynamic Bayesian networks. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 36(2):467–472, April 2006.

[166] Y. Zhu, E. Song, J. Zhou, and Z. You. Optimal dimensionality reduction of sensor data in multisensor estimation fusion. *IEEE Transactions on Signal Processing*, 53(5):1631–1639, May 2005.