PASSIVE INDOOR LOCALIZATION USING VISIBLE LIGHT

PASSIVE INDOOR LOCALIZATION USING VISIBLE LIGHT

BY

KHAQAN MAJEED, B.Sc., M.Sc.

a thesis
submitted to the department of Electrical & Computer Engineering
and the school of graduate studies
of mcmaster university
in partial fulfilment of the requirements
for the degree of
Doctor of Philosophy

Doctor of Philosophy (2021) McMaster University
(Electrical & Computer Engineering) Hamilton, Ontario, Canada

TITLE:	Passive Indoor Localization using Visible Light
AUTHOR:	Khaqan Majeed B.Sc. (Electrical Engineering), University of Engineering and Technology (UET) Lahore, Lahore, Pakistan M.Sc. (Electrical Engineering), King Fahd University of Petroleum and Minerals (KFUPM), Dhahran, Saudi Arabia
SUPERVISOR:	Dr. Steve Hranilovic
NUMBER OF PAGES:	xxiii, 188

Lay Abstract

A passive indoor visible light positioning system is proposed where the user does not hold any device or sensor tags for localization. The proposed system employs existing luminaires in an indoor area and does channel sounding between luminare-receiver pairs in order to locate the user. The presence of user alters measured diffuse light reflections that can be related to their position. The proposed work is divided into single-bounce model and realistic room model. The single-bounce model considers one bounce of the reflected light rays, whereas the realistic room model considers multiple bounces along with furniture inside the room. Furthermore, a fall detection system is developed to classify state of the user i.e., upright or prone. The proposed passive localization system leverages existing lighting infrastructure in the indoor environment and can be used to monitor patients in a home or hospital environment while keeping their privacy intact.

Abstract

In this thesis a proof-of-concept of passive indoor localization system using visible light is proposed that does not require active participation of a user in the localization process. The user neither holds any device nor do they have any sensor tags attached to their body. The system can be implemented by employing existing lighting infrastructure that is used in visible light communication systems.

The sources and receivers can be arranged in any form in the room, but in this work they are considered co-located on the ceiling. This arrangement is advantageous since it reduces installation complexity and is most commonly used in indoor environments. The proposed approach measures impulse response (IR) between the source-receiver pairs in order to localize a localization object (LO) i.e., the user. The presence of the LO inside the room alters the IR measurements between the source-receiver pairs that can be related to its position. The changes in measured IRs are leveraged for position estimation. The proposed research work can be divided into two main parts as follows.

In the first part, a single-bounce reflection model of light rays is considered and the room contains only LO inside it. The fingerprinting method is used to estimate position of the LO and analytical expression of Cramér-Rao lower bound is derived on the positioning error. In the second part, a realistic room model using reasonable parameters and multi-order reflections is considered where the furniture is also placed inside the room. A deep learning framework is employed that learns changes in IRs corresponding to random locations of LO in the room in order to estimate its position from the unknown IR measurements. Furthermore, a fall detection system is developed that classifies upright or prone states of the LO from single set of IR measurements.

To my beloved Mom, Dad, and Siblings

Acknowledgements

In the Name of Allah - the Most Compassionate, Most Merciful.
All praise is for Allah - Lord of all worlds.

First of all, I would like to thank my thesis advisor Dr. Steve Hranilovic for his continuous support and guidance throughout my Ph.D. degree program. His guidance and feedback helped me extensively in order to complete my research work. The valuable experience I gained is definitely going to help me in my future endeavors.

Then I would like to thank my committee members Dr. Thia Kirubarajan and Dr. Shiva Kumar for providing me with the valuable feedback during the committee meetings. I would also like to thank my external examiner Dr. Maïté Brandt-Pearce for providing me with the valuable insights and suggestions about my research work.

I would like to thank Drs. James Reilly, Ratnasingham Tharmarasa, and Tim Davidson for providing useful insights in some parts of my research work. Furthermore, I would like to thank my friends/colleagues Oussama Haddad, Muhammad Alrabeiah, Abdallah Shawky, Eric Mahoney, and Hossein Eldeeb for fruitful discussions about certain topics.

I would like to express my gratitude to friendly professors and staff at ECE department for providing me with a conducive environment in order to pursue my Ph.D. degree program.

The four and a half years of my stay at McMaster University helped me to have wonderful experience in terms of new friendships and outings with my friends. Many thanks to all those friends. I am affirmative that these memories are going to stay with me for a lifetime.

Finally, I would like to thank my Mom, Dad, and Siblings for their continuous moral support and unconditional love throughout my life, but especially during my Ph.D. journey. The frequent phone calls with them helped me to boost my morale and experience home-like feeling despite of living thousands of kilometers away from home.

Contents

List of Figures

List of Tables

Abbreviations

Abbreviations

AOA	Angle-of-arrival
AWGN	Additive white Gaussian noise
CAD	Computer-aided design
CRLB	Cramér-Rao lower bound
FC	Fully connected
FOV	Field-of-view
FIM	Fisher information matrix
FNN	Feed-forward neural network
GPS	Global positioning system
IoT	Internet of things
IR	Impulse response

LBS	Location based services
LED	Light-emitting diode
LiDAR	Light detection and ranging
LO	Localization object
LOS	Line-of-sight
MLE	Maximum likelihood estimation
MSE	Mean-squared error
NN	Neural Networks
OI	Object of interest
OIR	Object impulse response
PD	Photodiode or photodetector
ReLU	Rectified linear unit
RF	Radio frequency
RFID	Radio frequency identification device
RMSE	Root-mean-squared error
RSS	Received signal strength
SNR	Signal-to-noise ratio
STD	Standard deviation

TDOA	Time-difference-of-arrival
TO	Target object
TOA	Time-of-arrival
TOF	Time-of-flight
UWB	Ultra-wide band
VLC	Visible light communication
VLP	Visible light positioning

Chapter 1

Introduction

Positioning and navigation systems have undergone vast technological developments in the past few decades due to the development of sophisticated hardware e.g., mobile devices, inertial sensors, cameras, light sources, etc. This advancement has led to the deployment of positioning systems for variety of applications. Though there have been considerable advances to outdoor navigation, localization and navigation indoors is a difficult problem since the transmitted electromagnetic waves undergo attenuation and multi-order reflections due to the structure and layout of the area and materials used to coat walls, ceiling, floor, furniture, etc.

This chapter describes background of positioning in both outdoor and indoor scenarios. In particular, emphasis is made on indoor localization since the main contributions of the thesis fall under the category of indoor positioning. Indoor positioning has been sought out to a large extent lately due to high demand of location based services (LBS) in indoor environments. The applications related to outdoor and indoor positioning are also highlighted, especially the ones related to the proposed visible light-based passive indoor localization approach.

The organization of this chapter is as follows. The background of navigation is described in Sec. 1.1 with emphasis on outdoor scenarios and technology used for navigation. The challenges of using outdoor positioning technology in indoor environments is also discussed. Existing technologies used for indoor positioning are summarized in Sec. 1.2 while visible light positioning (VLP) is described in Sec. 1.3 where the concepts of active and passive localization are also explained. Sec. 1.4 gives an overview of existing localization techniques related to the VLP. The thesis layout and contributions are summarized in Sec. 1.5. Finally, the chapter is concluded in Sec. 1.6.

1.1 Positioning Background

The history of navigation dates back to 4000 BC when travelers used coast view [1] and constellations in night sky as reference points in order to reach their destinations. Digital technologies have recently been applied to provide accurate location information outdoors over the entire globe. In particular, outdoor navigation was revolutionized by the development of the global positioning system (GPS) [2] that became fully operational in 1995 [3]. The GPS consists of more than 24 satellites that orbit the Earth approximately 20,000 km away from its surface [3]. The GPS satellites in the Earth's orbit are arranged in order to cover a majority of its surface, however, the receiver on the Earth requires a subset of these satellites in its view for localization. In order to localize via GPS, the receiver must detect signals from at least four satellites in clear sky view and use trilateration methods (see Sec. 1.4.1). A recent method developed for smartphones using GPS provides accuracy around 5 m under open-sky [4]. However, GPS performance is degraded when signals are obstructed or weakened due to high-rise buildings or mountains. The advent of GPS led to many outdoor positioning applications e.g., navigating a user to a desired point of interest marked on the map, military applications, automobile navigation, and social networking to name a few [5].

In order to localize the user, there is a requirement of direct path i.e., line-of-sight (LOS) scenario between the GPS satellites and the receiver. The GPS is effective in outdoor environments as long as there is no blockage of signals between the receiver and satellites. The system does not work when the user carries the receiver inside the building or any closed indoor environment. This is due to the fact that the GPS signals become weak while entering the buildings due to large attenuation and fading [6]. These limitations of GPS have necessitated work in developing complementary indoor localization approaches in order to navigate users in indoor environments e.g., homes, offices, buildings, malls, airport, grocery stores, hospital, university campus buildings, etc.

1.2 Existing Indoor Localization Technologies

The need to provide LBS to people in indoor areas have led to the development of various methods and techniques that require mobile devices or sensor tags. Some recent surveys on methods and technologies used to deploy indoor positioning systems are described in [7, 8, 9, 10]. The choice of technology depends on many factors e.g., cost, accuracy, security and privacy, robustness, complexity, commercial availability and most importantly, the problem domain i.e., the type of localization application.

Existing technologies for indoor localization as described in [7, 8, 9, 10] can be grouped in the following manner: radio frequency (RF) (e.g., ultra-wide band (UWB), WiFi, radio frequency identification device (RFID), Bluetooth, Zigbee), inertial sensors in smart phones and mobile devices, audible sound or ultrasound, lasers or light detection and ranging (LiDAR), computer vision or camera, and visible light or infrared. The aforementioned technologies are briefly surveyed in the following.

1.2.1 Radio Frequency

In UWB-based techniques [11], a pulse of duration in the subnanosecond range is transmitted and detected at the receiver. The important characteristics of received signals such as time-of-arrival (TOA), time-difference-of-arrival (TDOA), angle-of-arrival (AOA), or signal strength (SS) are extracted and related to the position of receiver in order to localize the user. The large bandwidth of UWB signals enables them to pass through obstacles and furthermore, the received signals are less prone to interference and multipath effects [11]. Though UWB positioning methods provide high accuracy and high resolution [11], their usage is limited due to high cost of the equipment required [12].

In contrast to UWB, WiFi infrastructure is ubiquitously available in indoor areas [9] and is used to provide Internet connectivity to users. The WiFi infrastructure consists of WiFi access points (APs) that are arranged to cover majority of the indoor environment. Current mobile devices are capable of communicating with WiFi APs using RF signals and the communication between them is based on IEEE 802.11 standard [13]. The characteristics of signals received at the mobile device e.g., received signal strength (RSS) [7, 14, 15, 16], TOA, TDOA, AOA [17, 18, 19] can be measured and used to provide LBS to the user. Some techniques use a radio map of RSS values that is constructed in an offline phase [20]. The measurements obtained in the online phase are compared to the values stored in the map in order to estimate the position of the user. However, the radio map databases needs frequent updating due to temporal variations in the RSS values.

A similar set of measurements are used in Bluetooth and Zigbee networks [8]. However, these networks consume low power and cover a smaller range as compared to WiFi networks. A recent example of location awareness is introduced by Apple Inc. that uses Bluetooth to detect the presence of a mobile device in a certain region [21]. The system consists of a beacon that transmits signals periodically and a mobile device capable of receiving these signals which identifies whether it is in the vicinity of the signal transmitting beacon.

A similar category of location awareness techniques provide an approximate location of a user by using RFID tags [22, 23]. The RFID tag consists of an integrated circuit that is capable of storing information e.g., a unique code. This code is detected at an RFID reader and the measured RSS is used to approximate distance of the tag from the reader. The technique proposed in [23] enables location awareness for visually impaired people. The major challenge of RFID-based techniques is the requirement of infrastructure for localization. This is in contrast to WiFi infrastructure-based approaches which leverage existing wireless networks in the buildings in order to provide LBS to the users.

The choice of RF-based techniques depends on the type of localization application [9]. RF-based indoor localization systems are prone to security and privacy issues [10] since RF signals are not fully contained inside the room and can easily pass through walls of the room.

1.2.2 Inertial Sensors

Another class of indoor localization relies on measurements obtained from inertial sensors in smart phones [8, 24] e.g., proximity sensor, gyroscope, accelerometer, compass, camera, ambient light sensor, etc. An example of a typical measurement scenario is to count steps of a person walking in the area, which is then used to compute the distance covered by the person [24]. Similarly, these measurements can be used to detect the direction of movement or count the number of stairs climbed in order to compute the height traveled by the person [24]. Inertial sensor-based techniques are typically used in conjunction with WiFi infrastructure where measurements from both the inertial sensors and WiFi are used together to improve the position estimate of the user [8, 24].

The major challenge in using readings from inertial sensors is the accumulation of errors over time due to noisy sensors or random human behavior that adds bias to the sensor measurements [24]. This phenomenon is termed as dead reckoning and the errors occurred thereof in the measurements are minimized by using pedestrian dead reckoning algorithms [25].

1.2.3 Audible or Ultrasound Approaches

In audible source- or ultrasound-based techniques [26, 27, 28], acoustic sensors are used to measure the time-of-flight (TOF) of the received signals. Audible source-based systems use acoustic signals of frequency less than 20 kHz while ultrasound-based use ultrasound signals of frequency greater than 20 kHz. In order to localize, phase and frequency information are extracted from the received signals and are related to distance of a source from the receivers. The major challenge in using these techniques is the requirement of additional hardware and continuous power source (e.g., batteries) in order to obtain frequent measurement updates which eventually requires replacement or charging the batteries [10].

1.2.4 Laser or LiDAR Techniques

Laser- or LiDAR-based techniques [29, 30, 31] work on the principle that estimates range of a target from the LiDAR sensor using TOF between the transmitted and reflected signals, which is quite similar in principle to range estimation in radar systems [32]. LiDAR-based systems are extensively used in simultaneous localization and mapping (SLAM) techniques that produce a three-dimensional (3D) map of the indoor environment [33].

LiDAR systems work by transmitting a laser pulse to a target and measuring the reflected signals in order to extract features of the received signals e.g., TOF and wavelength. The extracted features are used to estimate the range of the target and are also labeled with respect to reference coordinates in order to build a 3D map of the environment. LiDAR-based systems have found applications in both indoor and outdoor environments. Modern vehicles are equipped with LiDAR scanners that help to build 3D map of a view ahead of the vehicle in order to notify drivers of potential warnings or assist in self-driving tasks [34]. Though LiDAR-based indoor localization systems provide estimation errors in centimeters [35], however, they are usually mounted on mobile robots or moving objects which requires the sensor to rotate or move in order to scan layout of the indoor area [35, 36].

1.2.5 Vision or Camera Methods

Vision- or camera-based techniques [37, 38, 39] use camera to acquire images of the indoor environment. The cameras can be installed in the indoor area [40] or on smart phones can be used to acquire images of the indoor area [41]. The captured images contain light beacons, where a beacon is defined as a light source that transmits a unique ID. An example of an indoor localization system that uses visible light communication (VLC) and sensors in smart phone can be found in a commercial system [42], where the proposed system provides geo-location of a mobile device in the indoor area. Such camera-based approaches pose high security risks since the map of an indoor area is stored on a phone or a cloud and is prone to unauthorized access [43]. The authors in [43] proposed a method that preserves privacy of the images gathered for localization by encoding 3D points in the indoor map to 3D lines that are used to conceal the original 3D map.

The visual sensor-based systems are less expensive as compared to the LiDAR-based systems (Sec. 1.2.4) and the future SLAM techniques have potential of information fusion from both LiDAR and visual sensors in order to provide improved 3D map of the indoor environment [44].

1.2.6 Visible Light or Infrared Systems

In visible light- or infrared-based localization systems, visible light or infrared sources are used to provide LBS to users. Lighting infrastructure in buildings is enabled to provide VLC to users in the area [45]. However, existing infrastructure can also be used to provide LBS to users without the need of requiring additional hardware resources [46]. The use of visible light-based localization systems proves reliable in terms of user privacy in indoor areas since the light rays are contained inside the area and do not travel through opaque walls [46]. The algorithms and techniques developed for visible light-based localization systems are often analogous to the ones in RF-based positioning [46, 47]. Visible light-based localization has several advantages over other technologies e.g., cost effectiveness, security, energy efficiency, ubiquitous availability of light sources in indoor areas, etc [46, 47]. Visible light-based localization is the main theme of the thesis and is discussed in detail in the remaining of this chapter and the thesis.

The technologies used for indoor positioning have their own pros and cons. The basic comparison of existing technologies is presented in Table 1.1.

1.3 Visible Light Positioning

The ubiquity of solid-state light-emitting diode (LED) light sources in indoor environments has revolutionized indoor communication systems. Consequently, VLC [45] systems have gained attraction recently due to simple infrastructure, low cost and high efficiency of LED light fixtures [48]. In order to provide current and future VLC applications, short-range optical wireless communications are standardized e.g., IEEE 802.15.7 [49] and IEEE 802.11bb [50].

Visible light positioning (VLP) systems can be considered an add-on to the VLC systems since they do not require the installation of extra hardware. They leverage existing modulated light sources from VLC networks and intensity detectors to infer user locations. The development of VLP [51, 46, 47, 52, 10, 53] systems has gained much momentum recently due to the high demand in indoor LBS e.g., guiding a user to the correct aisle in a grocery store, guiding people in shopping malls or airports,

Table 1.1: Comparison of Existing Indoor Localization Technologies

Technology	**Accuracy**	**Complexity and Challenges**
RF (UWB, WiFi, Zigbee, Bluetooth, RFID)	Few cm to few m (depending on the RF-based technology)	Increased complexity in hardware that require time synchronization, prone to security and privacy issues since RF signals travel through walls
Inertial Sensors	10's of cm to few m	Error accumulation due to dead reckoning, typically used in conjunction with WiFi
Audible or Ultrasound	Few m to around 20 m	Require continuous battery source, prone to noise pollution
Laser or LiDAR	Few cm	Require the sensor to move or rotate in order to scan the area, require time synchronization and generate large amounts of data for processing
Camera Systems	Order of few cm	Prone to user privacy issues since algorithms uses images of an indoor area
Visible Light or Infrared	Few mm to cm	Ambient light noise affects performance, time synchronization - small changes in time result in large changes in distance calculation

monitoring patients in a hospital environment, providing guided tours to the people in an art gallery, assisting students in university campus buildings, etc. [47, 51].

VLP systems have an added advantage in contrast to radio frequency (RF)-based localization systems in terms of user privacy and security. This is due to the fact that visible light is constrained within an indoor area and does not travel through opaque objects e.g., walls or adjacent rooms unlike RF signals which are prone to snooping [54]. VLP systems offer accuracy on the order of centimeters [46, 47] and can be broadly classified into active and passive localization systems based on the involvement of a user in the localization process and the measurement of specific signal parameters as well. Moreover, this classification can be application specific. Active and passive VLP scenarios are explained in the following.

1.3.1 Active VLP Scenario

In an *active* VLP scenario, a user is directly involved in the localization process e.g., a user equipped with a mobile device containing receiver or having sensor tags attached to their body. An example of active VLP scenario is shown in Fig. 1.1. The user holds receiver r that receives signals from sources s_i and s_j. Though only two sources are shown in the figure, in fact there could be receptions from more than two sources. Indoor areas typically contain multiple luminaires and they can be located on ceiling, walls, floor, mobile robots, etc. in order to provide LBS.

The scenario in Fig. 1.1 shows only direct LOS between the sources and receiver however, light rays can also reach the receiver after multi-order reflections. Typically such diffuse reflections are ignored and position estimates in active VLP methods rely on the energy received in LOS component of the received signals i.e., RSS. On the

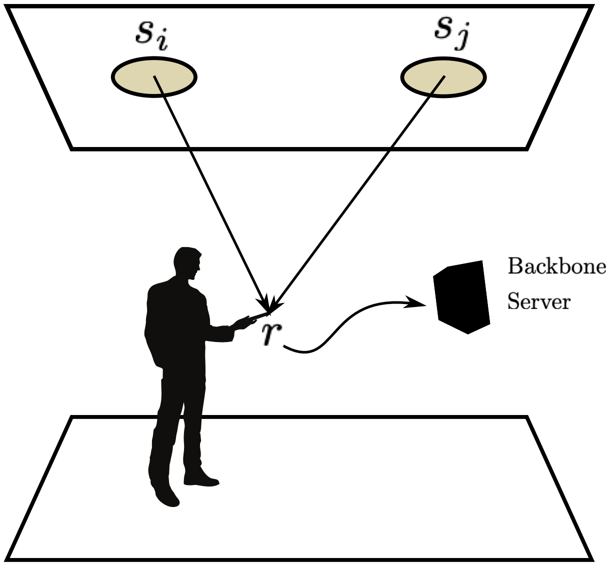


Figure 1.1: Active VLP scenario showing a LOS between the sources (s_i and s_j) and receiver (r). A user is holding a mobile device to receive signals, which are sent to a backbone server or can be processed at the mobile device in order to obtain their position estimate (clipart silhouette is reproduced from [55]).

other hand, TOA, TDOA, or AOA of the received signals can be measured in order to provide LBS as described in Sec. 1.4.1. The signals received at the mobile device are processed at the device or sent to a server at the back end in order to obtain position estimate of the user. The luminaires can be connected via power-over-Ethernet [56] and the mobile device needs to establish a connection with the network in order to obtain position estimate. A literature survey of different active VLP techniques is presented in Sec. 1.4.1.

1.3.2 Passive VLP Scenario

In *passive* VLP scenarios, a user is not directly involved in the localization process. An example case of passive localization is shown in Fig. 1.2, where the sources and receivers are affixed on the ceiling and are co-located. This arrangement of sources

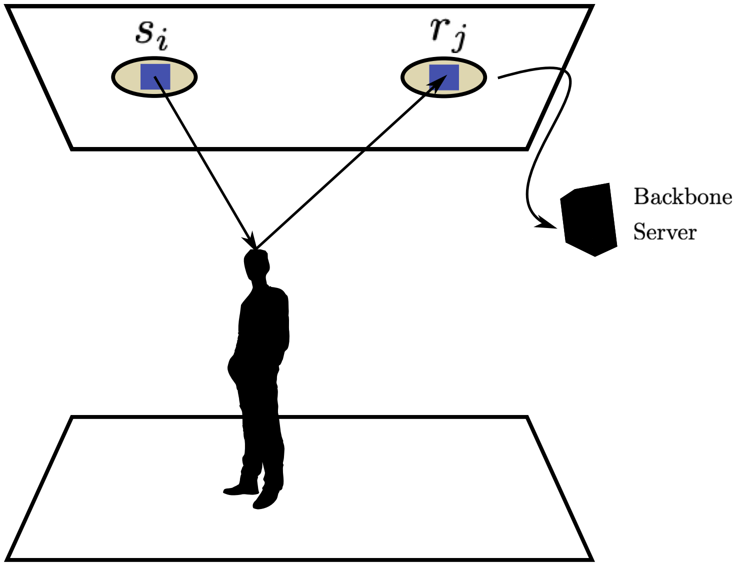


Figure 1.2: Passive VLP scenario showing sources and receivers co-located on the ceiling. An example ray is shown that originates from source s_i and is received at receiver r_j after bouncing off from the user's head. The signals received at the receivers are sent to a server at the back end in order to localize the user (clipart silhouette is reproduced from [55]).

and receivers is not strictly required, however, it is used in this research work since the indoor environments usually have symmetrical arrangement of luminaires on the ceiling.

In passive VLP, a user is detected in an indoor area by measuring reflected signals from the user or shadows formed due to their presence in the area [53]. The presence of a user in the area produces changes in signal measurements that are leveraged in localization. As observed from Fig. 1.2, this is important to note that the user neither holds any device in their hand nor do they have any sensor or photodetector (PD) attached to their body. Furthermore, it is not necessary to have a LOS between sources and receivers. A LOS scenario can also occur in passive VLP where the user can be detected in a region of blockage between the sources and receivers. Fig. 1.2

shows only first-order reflection, however, light rays can also reach the receivers after multi-order reflections. As an example, consider a light ray that emerges from source s_i and bounces off from the head of a person before it is received at the receiver r_j. The presence of a person inside the room produce changes in the received signals in contrast to when they are not present inside the room. Localization algorithms are developed that leverage these changes in order to estimate position of the user.

The recent abundance of mobile devices and smart devices and their further increase in future arose the concept of Internet-of-things (IoT), where these devices are capable of communicating with each other or a central server through the Internet. In contrast to active VLP, passive VLP relies on a backbone server in order to provide LBS since the user does not carry a mobile device. The concept of smart buildings show that luminaires can be controlled through power-over-Ethernet [56]. Ultimately, the luminaires in smart buildings can be connected to the IoT. This allows the control of luminaires from a central backbone network, which is advantageous to provide LBS in variety of applications ranging from casual services to healthcare domain to high risk indoor environments. A literature review of some passive VLP techniques is described in Sec. 1.4.2

1.4 Related Work in Visible Light Positioning

VLP uses a variety of signal measurements in order to develop localization algorithms. These algorithms typically depend on the type of signal characteristics and hardware infrastructure. The algorithms and techniques used in existing research work for VLP are described in the following.

1.4.1 Active Localization Techniques

Proximity-based Methods

Proximity-based methods are used in location awareness applications e.g., coarse localization stage in [57] or iBeacon [21] by Apple Inc. in case of RF-based localization. In proximity-based algorithms, the location of each luminaire is known and a unique code is assigned to it [46]. The receiver is detected when it receives signals from a luminaire under which it is present and the reported location is the area covered by the luminaire. An example of navigating a person indoors is proposed in [58] where the receiver detects a luminaire present in its vicinity by identifying the code transmitted by the luminaire.

Proximity-based techniques have low complexity, less accuracy and are cost effective as compared to other techniques [47]. These algorithms can be used in applications that do not require high accuracy (i.e., on order of cm) and can also be deployed in real time [46, 47]. The improvement in accuracy in these techniques occur at the expense of increasing the density of luminaires in an indoor area [47]. Some applications of proximity-based techniques are described in [51] e.g., guiding a person to a specific painting in art gallery.

Received Signal Strength-based Methods

RSS-based methods measure signal strength of the received signals at the receiver. The RSS values vary according to the propagation loss model that can be related to the position of the receiver. The various techniques and methods that employ RSS measurements for localization are described in the following.

The localization method proposed in [59] uses a fingerprinting approach where optical signals from multiple sources are collected to form a unique set of readings at a selected location. In order to develop a unique fingerprint at a discrete location, the LED light sources are modulated with distinct frequencies and the signals from these sources are detected at the receiver. The metric used for fingerprinting is the power spectral density (PSD) of each unique modulating frequency. The advantage of using PSD is that no synchronization is required between the light source and receiver. The modulated measurements are gathered at discrete locations in the indoor area during an offline phase in order to build up the fingerprinting map. In the online phase, the PSDs extracted from measurements with unknown position are then compared with the readings in the fingerprinting map. The reading with minimum error in the map provides position estimate of the user. The best reported accuracy of the fingerprinting method in [59] is around 15 cm at grid step size of 10 cm.

The method in [60] makes use of RSS to estimate location of the user. A Lambertian source model is used to relate distance to power measurements at the receiver. The position is then estimated using trilateration or multilateration. The authors in [61] provide a method for simultaneous indoor localization and high data rate VLC system. They employ discrete multi-tone (DMT) [62] modulation for positioning and use RSS-based ranging. Furthermore, a theoretical bound is also derived for estimation error.

The authors in [63] propose a localization method that uses received signal strength indication (RSSI) and time division multiplexing of the received signals in order to identify sources. The IDs of LED light sources are embedded in the control signal used for VLC and the receiver is mounted on a robot that moves in the indoor area. Similarly, the authors in [64] use RSSI for localization and also derive Cramér-Rao lower bound (CRLB) on the positioning error. The proposed method requires at least three sources and uses triangulation to estimate the position of the receiver.

The authors in [37] propose a localization method that uses LED sources as beacons. The LED beacons transmit light patterns that are detected at a digital camera equipped in the mobile device. The LED light source is partitioned into different parts in order to transmit a unique pattern, which is identified at the camera to estimate position of the user. A similar type of method that employs LED beacons is proposed in [39], however, the location of LED source is transmitted instead of physical partitioning of its surface as proposed in [37].

The authors in [65] propose a localization system that uses signal transmission in an up-link fashion i.e., a source located on the floor and receivers on the ceiling. A fingerprinting map is constructed in the offline phase where each fingerprint consists of a LOS component, diffuse component and time difference between them. The measurements with unknown position are obtained in the online phase and compared with the readings in a fingerprinting map in order to estimate the position of the user.

The authors in [66] use inertial sensors and ambient light sensor in smart phones in order to localize a user. The method uses the changing intensity values to infer the speed of motion of the user, however, it does require a map of the lighting in an area.

The accuracy of RSS-based VLP systems is reported as low as several centimeters, however, the accuracy is relatively lower around corners or edges of the room [46, 47]. Moreover, transmitted signals may undergo interference if adjacent rooms have glass walls or the performance may get worse if the room has ambient light [47].

Trilateration- or Triangulation-based Methods

Trilateration-based methods estimate the position of a receiver using its distance from the sources. The distance estimates can be obtained by using TOA or TDOA of the received signals and speed of light which is a constant [46, 47]. On the other hand, triangulation-based methods measure the AOA of the received signals in order to estimate position of the receiver [46, 47]. These techniques typically require received signals from at least three sources in order to estimate location of the user. The working principle in these techniques is similar to GPS outdoors where the location of the receiver is estimated using trilateration method.

The method proposed in [67] makes use of an aperture-based receiver for localization, which is similar to the one used in [68]. The aperture-based receiver consists of multiple PDs with apertures in front of them at a fixed distance. The apertures lie in a plane that is parallel to the PDs and their diameter is chosen larger than the wavelength of light. The light rays reach the PDs only through their apertures. The authors use a set of sinusoids as reference signals and correlate them with optical signals transmitted by LED light sources in order to distinguish between the sources. Furthermore, the CRLB is derived on location estimate of the user in order to evaluate performance of the system.

The method in [69] uses TOA-based ranging in order to provide a user's position estimate along with the derivation of CRLB. The method in [70] performs localization using TDOA, where LED light sources are assigned to unique frequency IDs. The light sources are assumed Lambertian and the optical channel is modeled as scaled and delayed Dirac delta function [45]. The TDOA is estimated at the receiver by measuring phase difference from three light sources. The estimated TDOAs are then used to localize the receiver.

The method proposed in [71] measures AOA at the receiver from multiple light sources. The LEDs are arranged as a grid on the ceiling and the receiver contains three PDs arranged in an orthogonal fashion. Such arrangement of PDs provides angular measurements at the receiver that helps to estimate the position of the receiver. A similar type of positioning algorithm is proposed in [72], where the receiver also consists of multiple PDs. The orientation of the PDs is first estimated in receiver's coordinate system. This information is then translated to the physical coordinate system and the position of the receiver is estimated along with its orientation.

The method in [73] provides position estimate of the user in 3D by combining RSS and AOA of the received signals. A single LED light source is considered in an indoor area, which is assumed to follow a Lambertian radiation pattern. The incident angle of the rays detected at the receiver are measured by an image sensor. The measured RSS and AOA then help to estimate the position of the receiver. Similarly, the work in [74] proposes a hybrid RSS/AOA technique, which makes use of signal strength, elevation and azimuthal angles in order to estimate position of the user. The proposed method consists of coarse localization and fine localization stages. In the coarse localization stage, an approximate location of the receiver is estimated using RSS at the receiver. In the fine localization stage, the position estimate is improved by triangulation.

The authors in [75] propose a localization approach which uses a novel luminaire topology to provide angular information directly to a single element detector (like an ambient light sensor on a mobile phone).

The reported accuracy of trilateration or triangulation-based techniques can be as low as few millimeters [47]. However, these techniques have an additional requirement that the sources and receivers are required to be time synced with each other. The geometric analysis of the indoor area and requirement of additional hardware make these systems complex in contrast to RSS-based methods [47].

1.4.2 Passive Localization Techniques

In Sec. 1.4.1 for active localization approaches the user is directly involved in the localization process i.e., transmitting received signals from multiple sources to a central server or processing the signals at a mobile device in order to obtain their position estimate. In contrast, passive localization approaches do not require the explicit actions of a user to provide positioning information. A summary of passive localization methods is described in [53], where these methods mainly use a unique arrangement of sources and receiver and/or additional hardware resources. Following describes some passive localization techniques.

The authors in [76, 77] propose a passive localization scenario in outdoor environment, where a bar code is used to assign an ID to the object that is similar to the one in [78]. The bar code is decoded at the receiver after sensing the reflected signals from the object. The decoded code and received power in reflected measurements help to identify and locate the object. The reported accuracy is on the order of several centimeters [77].

The authors in [79] study the blockage of signals between the sources and receivers in order to estimate location of an object. The user devices are located on the floor and transceivers on the ceiling. The proposed system offers a predicted accuracy on the order of few centimeters. However, performance is greatly affected due to multi-order reflections since the power in non-LOS components may be detected as the power in LOS component. Secondly, the proposed system requires sources located on the floor that results in modifying the lighting infrastructure in the indoor environment since the typical arrangement of luminaires is on the ceiling in the area. In [80], the proposed system model is similar to [79] and the authors consider multi-order reflections while simulating the channel model. The performance bounds are computed on fingerprinting-based localization and a nearest neighbor fingerprinting approach is used to evaluate tightness of the derived bound. The reported error is near 5 cm using 4 PDs in the area.

The method in [81] uses receivers that are attached to a wall instead of floor or ceiling. The changes in RSS at the receivers are sensed due to the presence of an object and, therefore, are used to estimate the location of the object. The authors in [82] use shadows of light formed due to humans in the indoor area in order to detect them and construct their 3D postures.

The authors in [83] makes use of time slots to transmit signals from different sources in order to identify them at the receiver. The proposed system senses shadows formed by an object in order to identify its occupancy in the room. The position of the object is estimated by measuring changes in the power detected at the receiver. The reported median error is 0.89 m in a room of dimensions 7.5m×6m×2.74m containing a table in the center. A similar type of methodology is proposed in [84] that estimates location of the object and occupancy in the indoor environment by sensing changes in voltage measurements detected at the receiver. The signals are transmitted in time division multiplexed format and are identified at the receivers. Moreover, the algorithm requires positions of light sources to be known in advance.

The method in [85] also uses shadows cast on the floor in order to estimate location of an object. However, in contrast to [83, 84], the floor contains PDs equipped with RFID tags in order to power the PDs. A user is sensed under the luminaire by detecting blockage of light between the luminaire and PD.

Existing passive localization techniques [53] either require the modulation of light sources and using polarizable materials in the indoor area (e.g., [86]) or modification of lighting infrastructure in the indoor area along with positioning of receivers (e.g., receivers on walls in [81]). Moreover, the techniques use certain characteristics of the received signals e.g., voltage measurements, RSS, or shadow-based RSS that offer challenges in positioning in case of multi-order reflections.

The aforementioned arrangement of luminaires and receivers makes the system deployment application specific. One of the main goals in designing passive indoor localization systems is to use existing lighting infrastructure with minimal modifications in order to enable deployment of localization systems feasible in multiple indoor scenarios and environments. Secondly, the objective is to relieve user from carrying a mobile device or sensor.

1.5 Thesis Contributions and Outline

This section summarizes the main contributions and general outline of the thesis. The thesis is written in "*sandwich-style*" format, where the chapters correspond to individual papers and journal articles. The papers have been modified slightly in order to follow the thesis format, however, the main content in the papers is unchanged.

The research work presented in the papers and consequently, restated in the thesis is completed solely by Khaqan Majeed under the supervision of Dr. Steve Hranilovic. The manuscripts and thesis have been improved in clarity and presentation through feedback received from Dr. Hranilovic. The acknowledgments are made in some papers and the thesis in regards to useful discussions that were carried out with the concerned persons.

The research work presented in this thesis arises from published/submitted papers or articles, which are listed in the following.

1. [87] **K. Majeed** and S. Hranilovic, "Passive Indoor Localization for Visible Light Communication Systems," *2018 IEEE Global Communications Conference (GLOBECOM)*, Abu Dhabi, United Arab Emirates, 2018, pp. 1-6, doi: 10.1109/GLOCOM.2018.8647875.

2. [88] **K. Majeed** and S. Hranilovic, "Performance Bounds on Passive Indoor Positioning Using Visible Light," in *Journal of Lightwave Technology*, vol. 38, no. 8, pp. 2190-2200, 15 April 15, 2020, doi: 10.1109/JLT.2020.2966365.

3. [89] **K. Majeed** and S. Hranilovic, "Passive Indoor Visible Light Positioning System using Deep Learning," in *IEEE Internet of Things Journal*, doi: 10.1109/JIOT.2021.3072201.

4. [90] **K. Majeed** and S. Hranilovic, "Passive Indoor Visible Light-based Fall Detection using Neural Networks," *under preparation.*

5. [91] **K. Majeed** and S. Hranilovic, "Passive Positioning using Visible Light Systems," invited paper in *The Optical Networking and Communication Conference & Exhibition (OFC) 2021, OSA.*

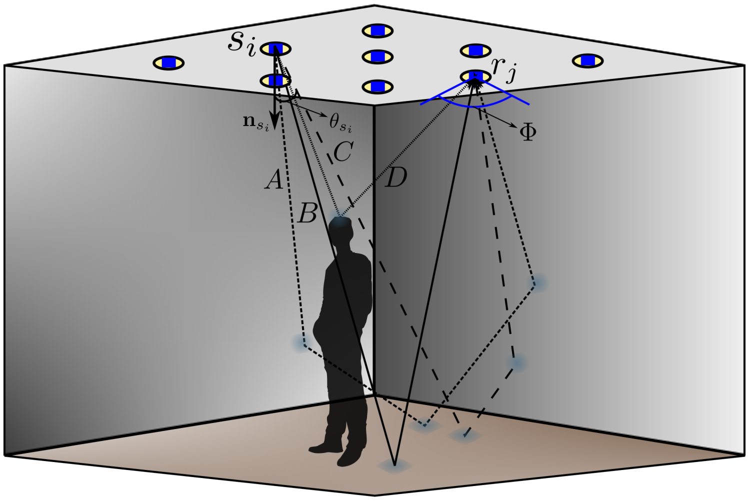


Figure 1.3: Passive scenario showing co-located light sources and receivers on the ceiling. The rays undergo multi-order reflections before being detected at the receiver. A clipart silhouette is reproduced from [55] (Figure is taken from our paper [89]).

1.5.1 Summary of Contributions

The passive techniques surveyed in Sec. 1.4.2 require rearrangement of luminaires and receivers in an indoor area in order to provide LBS. Some passive techniques described in [53] require users to have sensor tags attached to their body while others require installation of RFID tags on the floor. Some techniques rely on only RSS measurements or RSS-based shadows in the area, which in case of multi-order reflections affects performance of the system.

The proposed method relieves user from carrying any mobile device or sensor tags and provides a proof-of-concept of passive indoor localization system that leverages existing lighting infrastructure. The proposed passive indoor localization paradigm is described in the following.

Proposed Passive VLP Paradigm

The proposed passive localization scenario is shown in Fig. 1.3 where the sources and receivers are co-located on the ceiling. The terms localization object (LO), object of interest (OI), and target object (TO) are used interchangeably in the thesis to represent the same object i.e., a user whose position estimate is required. The LO is placed inside the room and impulse response (IR) measurements are obtained between all source-receiver pairs. The presence of LO inside the room produce changes in IRs that can be related to its position. The fundamental notion behind the proposed passive approach is to estimate position of the LO using the changes observed in the IRs that occur due to the presence of LO inside the indoor area.

In order to measure IRs between the source-receiver pairs, each source is turned on at a time and IR acquisitions are obtained at all the receivers. This process is repeated for all the sources. The time extent of an IR can be upper bounded by the geometry of the indoor environment and the location of sources and receivers. The length of each IR is usually on the order of 10's of ns and thus all IRs can be accumulated within fraction of a second.

It is important to note that the IR measurements are affected by limited rise time of the luminaires and receivers. In the IR acquisition process it is assumed that the employed devices have sufficient bandwidth ratings in order to capture details of the indoor environment corresponding to room furnishings, especially the peaks occurring in IR due to the presence of the LO inside the room.

The proposed research results are divided into two main parts based on the type of room model used in the localization process. The first part employs a single-bounce model where the room is considered empty with first-order reflections only.

The single-bounce model is tractable and permits for the derivation of theoretical results on performance. The second part uses a realistic room model with multi-order reflections and furniture inside the room. The sources are modeled as having wavelength-dependent flux variation and the surfaces of different entities in the room are coated with real coating materials that have wavelength-dependent reflectivity values. The realistic room model provides a proof-of-concept of passive VLP to understand performance in a realistic scenario since the actual room contains furniture and the IR measurements include multi-order reflections.

1.5.2 Single-bounce Model

As described in Sec. 1.5.1, the single-bounce model considers only first-order reflections in an empty room in order to model the IRs. The first-order reflections help to derive the theoretical performance of the proposed passive indoor VLP system. The proposed approach considers existing luminaire infrastructure. The changes in IR measurements between source-receiver pairs due to the presence of LO in the room correspond to its location. The IR measurements, in contrast to RSS-based measurements [53], inherently contain time information along with amplitude of a peak that corresponds to the location of LO in the room. The peak occurring in IR corresponding to LO is dominant in first-order reflection from the LO as compared to higher order reflections and furthermore, the amplitude and time delay of the peak depend on location and height of the LO. Chapters 2 and 3 employ single-bounce model and the contributions therein are described in the following.

Chapter 2 - Passive Indoor Localization for Visible Light Communication Systems

The research work described in Chapter 2 appeared in the following conference proceeding.

- [87] **K. Majeed** and S. Hranilovic, "Passive Indoor Localization for Visible Light Communication Systems," *2018 IEEE Global Communications Conference (GLOBECOM)*, Abu Dhabi, United Arab Emirates, 2018, pp. 1-6, doi: 10.1109/GLOCOM.2018.8647875.

Chapter 2 describes fingerprinting method, which is used to estimate position of the LO [87]. The proposed room scenario is based on single-bounce model where the room is considered empty along with first-order reflections only. The luminaires and receivers are affixed on the ceiling in a co-located fashion. A network of luminaires and receivers can be controlled from a central location at the back end as shown in Fig. 1.2. The proposed approach measures the IR between the source-receiver pairs. In order to represent the IR due to the LO only, IR differences are obtained by subtracting the IR of the empty room from the IR when the LO is present inside the room. The IR differences are modeled using an exponential integrating-sphere model [92], where the amplitude, time delay, and decay rate of the fitted model correspond to the location of the LO at a certain discrete positions in the room. The fitted model parameters corresponding to LO at discrete locations are stored in a fingerprinting map instead of storing complete waveforms in order to reduce storage requirements. It is important to note that the fingerprinting map is obtained in the offline phase.

In order to estimate the position of the LO at an unknown location, the IRs are measured in the online phase and then IR differences are obtained. The exponential integrating-sphere model [92] is fitted to the IR differences in a similar fashion to the readings in the fingerprinting map. The fitted model is then compared to the readings stored in the map. The closest point in the map using a minimum error criterion is reported as a position estimate of the LO. The system performance is evaluated by using fingerprinting map with different grid spacing and varying brightness of the luminaires.

Chapter 3 - Performance Bounds on Passive Indoor Positioning using Visible Light

The proposed research work in Chapter 3 appeared in the following journal paper.

- [88] **K. Majeed** and S. Hranilovic, "Performance Bounds on Passive Indoor Positioning Using Visible Light," in *Journal of Lightwave Technology*, vol. 38, no. 8, pp. 2190-2200, 15 April 15, 2020, doi: 10.1109/JLT.2020.2966365.

Chapter 3 describes the derivation of a CRLB on the positioning error of LO in an empty room that is modeled using single-bounce model [88]. The acquisition of IRs and IR differences are obtained in a similar fashion as described in Sec. 1.5.2. The IR differences are zero-clipped to remove negative part in the difference in order to fit the exponential integrating-sphere model [92]. The parameters of the fitted model i.e., amplitude, time delay, and decay rate can be related to the position of LO inside the room as described earlier.

In order to derive analytical expression of the CRLB, the Fisher information matrix is obtained that consists of partial derivatives with respect to individual Cartesian coordinates of the position of LO. Furthermore, a maximum likelihood estimator is used to estimate position of the LO and evaluate the tightness of the derived bound.

The variation of CRLB over the spatial domain of the room is calculated in order to observe the performance variation over the room since some receivers are unable to see the LO in certain regions of the room due to limitation in their field-of-view. The effect of brightness of the luminaires and the number of source-receiver pairs used in order to derive the bound are also studied.

1.5.3 Realistic Room Model

Though the simpler single-bounce model is tractable for theoretical study, in order to characterize the performance of passive VLP a more complete model is required. As described in Sec. 1.5.1, realistic room model considers multi-order reflections and all surfaces and entities in the room are simulated with coatings which have reflectivity values measured for real materials [93, 94, 95]. The realistic room model provides a proof-of-concept of passive VLP in order to characterize the performance in a realistic scenario. Chapters 4 and 5 employ realistic room model and the contributions therein are described in the following.

Chapter 4 - Passive Indoor Visible Light Positioning System using Deep Learning

The research work presented in Chapter 4 is accepted for publication in the following journal.

- [89] **K. Majeed** and S. Hranilovic, "Passive Indoor Visible Light Positioning System using Deep Learning," in *IEEE Internet of Things Journal*, doi: 10.1109/JIOT.2021.3072201.

Unlike Chapters 2 and 3, the indoor scenario used in Chapter 4 is based on realistic room model. The luminares and receivers do not have LOS between them since they are co-located on the ceiling as described in Sec. 1.5.2. However, the luminaires are assumed to have wavelength-dependent characteristics. The coating materials used to represent the optical properties of different surfaces and objects in the room are also assumed to have wavelength-dependent reflectivity values. Furthermore, the room contains furniture and the IRs that are measured in the same way as described in Sec. 1.5.1, however, the IRs in this case contain multi-order reflections. A deep learning framework composed of a feed-forward neural network architecture [96] is employed in order to learn the relationship between IRs and position of LO since the IRs contain multi-order reflections and the room contains furniture which is in contrast to the single-bounce model described in Sec. 1.5.2. Multiple sets of IRs are obtained when the LO is present at randomly selected positions in the room, where each set of IRs correspond to measurements from all source-receiver pairs in the room. The presence of a LO at a certain location produce changes in IRs, which are crucial in the position learning process.

The gathered data are divided into training and validation data sets and preprocessed before feeding them to the deep learning framework. It is important to note that only training data are used to train the network, whereas the validation data are used only to evaluate performance of the network. The trained network is then used to estimate position of the LO located at unknown location in the room by using a set of IR measurements at that location. The reported root-mean-square error is around 80 cm by using sets of IR measurements at as low as 10 randomly selected training positions in the room.

Chapter 5 - Passive Indoor Visible Light-based Fall Detection using Neural Networks

The research work in Chapter 5 is completed and the paper is under preparation.

- [90] **K. Majeed** and S. Hranilovic, "Passive Indoor Visible Light-based Fall Detection using Neural Networks," *under preparation*.

Chapter 5 describes a passive fall detection system based on visible light that detects human fall in a room. The room scenario employs realistic room model similar to the one described in Chapter 4. The measurement process of IRs is the same as described in Sec. 1.5.1. Feed-forward neural networks [96] are used in order to classify the state of the user i.e., upright or prone. In order to gather data (training and validation) to train the network, the LO is placed at random positions in the room and assumed to have either of the following states: upright or prone. The prone state of the LO is modeled as lying on the floor along either one of the following directions: horizontal (x-axis) or vertical (y-axis).

The set of IR measurements are obtained corresponding to upright or prone states of the LO and are labeled into two classes. The labeled data are divided into training and validation sets and preprocessed before feeding them to the neural network classifier. The trained network is then employed to estimate state of the LO using a set of acquired IR measurements in order to detect the fall. The sensitivity of the proposed technique is evaluated by considering an additional state (tilted) that models the LO to lean towards one of the following directions: north, east, south, and west. This state is labeled with the same class as the upright state in order to differentiate it from the prone state. The observation of results show that the algorithm maintains similar performance while classifying the prone state from upright or tilted sates. The proposed system shows accuracy greater than 97% in predicting upright and prone states of the user even when the measurements corresponding to tilted state are included in the validation data (training data is comprised of only upright and prone states).

1.5.4 Chapter 6 - Conclusions and Future work

The conclusions of the research work proposed in the thesis are summarized in Chapter 6. Furthermore, possible directions for continuation of the research work are also described in this chapter.

1.6 Conclusions

A brief overview of outdoor navigation systems is presented since outdoor navigation has evolved and matured over time, especially during the recent past. The global positioning system (GPS) is widely used in outdoor environments in order to provide location based services (LBS). The GPS signals undergo high attenuation while passing through buildings. Therefore, the need arises to seek other technologies for localization in indoor environments. Existing technologies used for indoor localization are radio frequency (UWB, WiFi, RFID, Bluetooth, Zigbee), inertial sensors in mobile devices, audible or ultrasound, LiDAR, camera, visible light or infrared.

The choice of technology typically depends on the domain of application. VLP offers a cost effective solution due to the use of light-emitting diode sources and preserves user privacy because the light rays are contained within the room. Moreover, future luminaire fixtures have the capability of being controlled through internet-of-things devices, which enables a convenient means to provide a backbone connection to a server for a VLP system.

VLP methods can be broadly divided into active and passive localization scenarios. The active scenario involves a user directly in the localization process which is in contrast to passive scenario. Active localization methods use different signal characteristics e.g., received signal strength, time-of-arrival, time-difference-of-arrival, angle-of-arrival, or their hybrid combination. Nevertheless, the user is required to carry some device or sensor in order to access the LBS.

The proposed research work presents a proof-of-concept of passive indoor localization system that is based on visible light. The work is divided into two parts based on the room model. In the first part, a single-bounce model is used where the room is considered empty and impulse responses (IRs) contain first-order reflections only. The single-bounce model helps to characterize the performance of the passive localization systems theoretically. A position estimate of the localization object (LO) is obtained by using fingerprinting method. Furthermore, analytical expression of Cramér-Rao lower bound is derived on the positioning error spatially over the room. In the second part, a realistic room model is considered with furniture inside it and the IRs contain multi-order reflections. A deep learning framework is employed in order to estimate position of the LO from the gathered data that is obtained by placing the LO at uniformly distributed locations in the room. Furthermore, a fall detection system is developed that detects state of the LO i.e., upright or prone. Feed-forward neural networks are employed in order to classify the state of LO from a set of IR measurements.

The proposed passive indoor localization techniques can be used in variety of applications where the user is not required to carry any mobile device with them e.g., monitoring people in private indoor environments, locating workers in high-risk indoor areas, monitoring seniors at home or hospitals in order to provide timely healthcare, etc.

Chapter 2

Passive Indoor Localization for Visible Light Communication Systems

IEEE Copyright Notice

The research work presented in this chapter appeared in the following conference paper [87].

- **K. Majeed** and S. Hranilovic, "Passive Indoor Localization for Visible Light Communication Systems," *2018 IEEE Global Communications Conference (GLOBECOM)*, Abu Dhabi, United Arab Emirates, 2018, pp. 1-6, doi: 10.1109/GLOCOM.2018.8647875.

The paper is reproduced in this chapter with minor modifications in formatting in order to follow the thesis format i.e., setting of equations, equation numbers, citations, figures, etc.

In this chapter, single-bounce model of the room is used where first-order reflections are considered only and source-receiver pairs are co-located on the ceiling. A fingerprinting method is used to localize a localization object (LO) i.e., a person. In order to build up the fingerprinting map, the LO is placed at predefined grid points and impulse response (IR) measurements are obtained between all source-receiver pairs. The IR due to LO only is obtained by calculating IR difference, where the difference is obtained by subtracting the IR of empty room from the IR when the LO is present inside the room. An exponential integrating-sphere model [92] is fitted to the IR differences. The fitted model provides three important parameters namely amplitude, time delay and decay rate of the exponential. The fingerprinting map contains the parameters of the exponential model at discrete locations. This greatly reduces storage costs as compared to storing the complete IR waveform at each location. The IR differences in the online phase are then compared to the ones stored in the fingerprinting map in order to estimate the position of LO.

2.1 Abstract

In this paper we present a new methodology for indoor localization using visible light emissions which does not require users to actively participate in the process. Our *passive positioning* approach utilizes a network of luminaires and receivers on the ceiling and measures the impulse responses (IRs) between each source-receiver pair. This channel sounding approach leverages changes in the measured IRs to localize an object in the room. Simulation results show that with a database of sampled IRs, the root mean squared (RMS) positioning error can be made on the order of 4 cm. In order to lower storage requirements, a simplified database based on fitting to an integrating sphere model of the IR yields localization performance with RMS error around 6 cm.

2.2 Introduction

The use of solid-state light-emitting diodes (LEDs) for indoor illumination is nearly ubiquitous. The economic and energy-efficiency benefits of LEDs make them nearly ideal candidates for indoor lighting. With the proliferation of LED illumination, applications such as visible light communications (VLC) and visible light positioning (VLP) have become attractive. In particular, the use of visible light for indoor localization offers benefits of privacy and security since all light emissions are confined to a given room [51, 46, 97, 98].

Existing indoor VLP techniques rely on users to provide measurements of light, perhaps via a smartphone, to enable localization. These techniques can be broadly classified into the following categories.

In *proximity*-based techniques, each light source transmits a unique emission which can be identified at the receiver. The approximate location of the user is estimated using this unique ID which is detected at the receiver. In the context of navigation for the visually impaired, [99, 100] experimentally demonstrated a coarse localization approach using simple receivers carried by the users. *Scene analysis* techniques, such as those in [37, 39], use a camera as receiver and infer position by using luminaires as beacons. The identity of the luminaire is communicated through LED spatial patterns or temporal changes which are sensed by the digital camera receiver. The third category are *triangulation* techniques, which rely on range, direction of the received signal, or a combination of both. In [101], each luminaire is assumed to identify itself with a unique temporal sequence. A position estimate is achieved assuming a known Lambertian emission pattern and location for each luminaire. User measurements of received signal strength are fit to a simplified propagation model. In [102] a hexagonal arrangement of LED sources is exploited to reduce coverage holes as well as the areas where coverage overlaps. Time delay-of-arrival information in conjunction with an Extended Kalman Filter is used to localize the receiver and improve position estimation in the indoor area. The method in [71] relies on angle-of-arrival measurements from the LED sources. The LEDs are arranged in a grid on the ceiling while a complex photo-receiver consisting of three photo diodes (PDs) arranged in a corner cube are used to measure arrival angle.

In recent work, [65] considers users with infrared beacons which are measured by sensors in a room. Both line-of-sight (LOS) and non-LOS components are tabulated and used for localization. Unlike earlier approaches, [65] localizes using an uplink channel rather than a downlink VLC channel.

In this paper, we propose a fundamentally new approach to indoor VLP termed *passive localization* in which the user does not explicitly participate in the localization process. Such type of localization provides convenience at the user's side since the user does not need to carry a receiver. The proposed technique promises to be useful in localizing elderly people at home since the timely help can be provided to them if they fall down on the floor and get injured. Luminaires are assumed to be distributed in a room in order to provide sufficient illumination. Optical receivers are also assumed to be present on the ceiling and can be potentially co-located with the LED sources. All sources are assumed to be connected via a network, such as power-over-ethernet or power line communications, in order to synchronize and communicate measurements to a central server. Localization is performed based on repeated measurements of the impulse responses (IRs) between all sources and receivers. Changes in the measured IRs are then used to infer the presence and location of an object or person in the room. This process is similar to channel sounding in the context of underwater acoustic channels, where dedicated signals are transmitted through the channel and IRs are measured [103]. Notice that in contrast to existing *active* VLP approaches, our positioning technique does not require the use of specialized beacons, luminaire arrangements or user interaction to enable the localization process at the expense of additional complexity in networking the luminaires.

The remaining parts of the paper are organized as follows. The system model for the proposed passive localization technique is explained in Section 2.3. The localization algorithm is presented in Section 2.4 and the performance of proposed method is evaluated in Section 2.5. Finally, the paper concludes in Section 2.6.

2.3 System Model

2.3.1 Problem Setup

Consider the localization scenario depicted in Figure 2.1 where the goal is to determine the location of an object (in this case a cuboid). Luminaires and receivers are located on the ceiling at known positions. Define $\mathcal{S} = \{s_i : i = 1, 2, \cdots, N_s\}$ as the collection of luminaires within the room and similarly, $\mathcal{R} = \{r_j : j = 1, 2, \cdots, N_r\}$ as the collection of receivers, where N_s and N_r are the total number of sources (i.e., luminaires) and receivers respectively.

Light emitted from each source is reflected from the floor, ceiling, walls and the object and is ultimately detected at each of the receivers in $\mathcal{R}$. A key assumption in this work is that the IR $h_{s_i r_j}(t)$ for any $s_i \in \mathcal{S}$ and $r_j \in \mathcal{R}$ can be measured independently. This can be accomplished by scheduling localization intervals via a network between the luminaires to ensure a single source is turned on individually and measurements of IR are collected. Given that time extent of $h_{s_i r_j}(t)$ is on the order of 10's of ns, the frequency of these measurements can be made sufficiently fast to make them imperceptible. It is further assumed that a database of $h_{s_i r_j}(t)$ between sources-and-receivers is measured and known at several points in the room with and without the presence of the object. The position of an object in the room can be determined by measuring the set of IRs between all sources-and-receivers and using the database. Notice that the object o, in Figure 2.1, does not participate explicitly in localization and hence the term *passive* localization is adopted. This IR measurement process is similar to *channel sounding* used in acoustic communication, where the channel IR is measured using probe-signals transmitted within the environment.

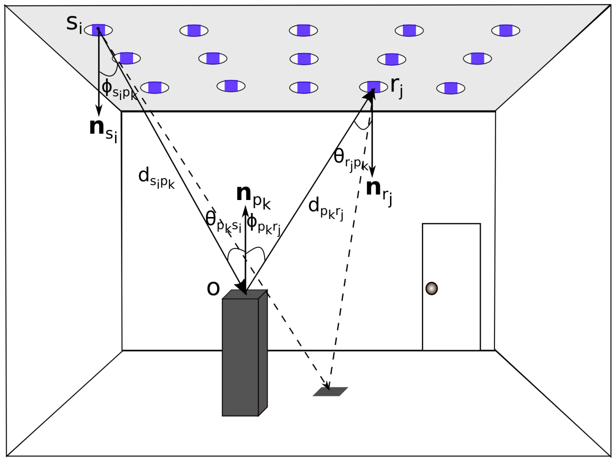


Figure 2.1: A room showing co-located light sources and receivers affixed on the ceiling with presence of an object (i.e., a cuboid). The light ray from source s_i is reflected from a patch on the object or floor (or walls) and received at receiver r_j.

2.3.2 Optical IR for Indoor Channels

For N_s luminaires and N_r receivers, define the collection of IRs between each source-receiver pair in an empty room (i.e., without the object) by set $\mathcal{H}_e$ as

$$\mathcal{H}_e = \left\{ \begin{array}{c} \{h_{s_1 r_1,e}(t), h_{s_1 r_2,e}(t), \cdots, h_{s_1 r_{N_r},e}(t)\} \\ \{h_{s_2 r_1,e}(t), h_{s_2 r_2,e}(t), \cdots, h_{s_2 r_{N_r},e}(t)\} \\ \vdots \\ \{h_{s_{N_s} r_1,e}(t), h_{s_{N_s} r_2,e}(t), \cdots, h_{s_{N_s} r_{N_r},e}(t)\} \end{array} \right\} \tag{2.3.1}$$

In practice, $\mathcal{H}_e$ could be measured by turning each source on one at a time and performing measurements which are stored in a database *a priori*. These IRs do not change appreciably in time and can be accurately estimated using many repeated measurements in the room.

Consider now that an object o is present inside the room as shown in Figure 2.1. In a similar fashion, define the collection of IRs when the object is located at position

(x, y, z) as

$$\mathcal{H}_o(t; x, y, z) = \left\{ \begin{array}{c} \left\{h_{s_1 r_1,o}(t; x, y, z), \cdots, h_{s_1 r_{N_r},o}(t; x, y, z)\right\} \\ \left\{h_{s_2 r_1,o}(t; x, y, z), \cdots, h_{s_2 r_{N_r},o}(t; x, y, z)\right\} \\ \vdots \\ \left\{h_{s_{N_s} r_1,o}(t; x, y, z), \cdots, h_{s_{N_s} r_{N_r},o}(t; x, y, z)\right\} \end{array} \right\} \tag{2.3.2}$$

The measured $\mathcal{H}_o(t; x, y, z)$ depends on the location of object as well as the background. In order to estimate the IR due to the presence of object only (i.e., without ceiling, floor, walls and background elements) consider the approximation

$$h_{s_i r_j,d}(t; x, y, z) = h_{s_i r_j,o}(t; x, y, z) - h_{s_i r_j,e}(t) \tag{2.3.3}$$

where $h_{s_i r_j,d}(t; x, y, z)$ is termed an *IR difference* and is difference between the IR when the object is located at (x, y, z) and the empty room. Define $\mathcal{H}_d(t; x, y, z)$ as

$$\mathcal{H}_d(t; x, y, z) = \left\{ \begin{array}{c} \left\{h_{s_1 r_1,d}(t; x, y, z), \cdots, h_{s_1 r_{N_r},d}(t; x, y, z)\right\} \\ \left\{h_{s_2 r_1,d}(t; x, y, z), \cdots, h_{s_2 r_{N_r},d}(t; x, y, z)\right\} \\ \vdots \\ \left\{h_{s_{N_s} r_1,d}(t; x, y, z), \cdots, h_{s_{N_s} r_{N_r},d}(t; x, y, z)\right\} \end{array} \right\}. \tag{2.3.4}$$

Notice that $\mathcal{H}_d(t; x, y, z)$ does *not* represent the collection of IRs of the object alone because of shadowing of the sources and receivers by the object. The location of the shadow depends on which light source and receiver is used as well as on the location of the object. Consider the example in Figure 2.2 which plots the DC gain of the IR from s_1 to r_6. There are two shadows present: shadow A arises due to

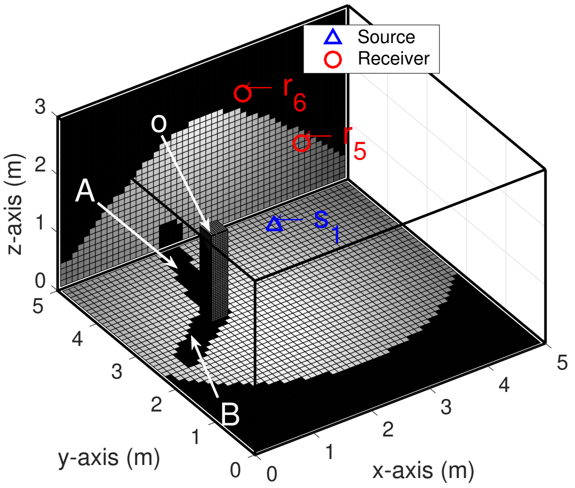


Figure 2.2: Plot of the DC gain of the IR $h_{s_1 r_{6,o}}(t)$. The object is located at $(1.5\,\mathrm{m}, 3.25\,\mathrm{m}, 1.5\,\mathrm{m})$. The source s_1 is located at $(1\,\mathrm{m}, 1\,\mathrm{m}, 3\,\mathrm{m})$ and receivers r_5 and r_6 are located at $(2.5\,\mathrm{m}, 2.5\,\mathrm{m}, 3\,\mathrm{m})$ and $(2.5\,\mathrm{m}, 4\,\mathrm{m}, 3\,\mathrm{m})$ respectively. ‘A’ denotes the shadow of object from s_1 while ‘B’ denotes the shadowing of receiver r_6 by the object.

the object shadowing s_1 and shadow B arises due to the obstruction of r_6 by o. In practice, however, the difference between the IRs in $\mathcal{H}_d\,(t; x, y, z)$ and the IRs due to the object alone is small due to the negligible amount of power neglected in the shadowed regions and is quantified in Section 2.5.

2.3.3 Database of IR Differences

The indoor area is assumed to be discretized at some regular spacing interval, e.g., at points (x_n, y_n, z_n) where $n = 1, 2, \cdots, N_x \times N_y \times N_z$. With an empty room, the dataset $\mathcal{H}_e$ is measured between all source-receiver pairs. In a calibration step, consider moving the object to each grid point where measurements of $\mathcal{H}_o\,(t; x, y, z)$ are made and the data set $\mathcal{H}_d\,(t; x, y, z)$ in (2.3.4) of IR differences is collected at each grid point. The resulting database is represented as $\{\mathcal{H}_d(t; x_n, y_n, z_n)\}$.

2.3.4 Extraction of IR Parameters

The approach in Section 2.3.3 requires a large database of sampled time waveforms representing IR differences to be stored. As an alternate approach, consider forming a smaller database which stores parameters extracted from the measured IR differences. Consider fitting all IR differences to the exponential integrating-sphere model for indoor diffuse optical wireless communications [104],

$$h(t) = \eta e^{-\frac{(t-t_1)}{\tau}} u(t - t_1) \tag{2.3.5}$$

which is parameterized by amplitude η, time delay t_1, delay spread τ and $u(t)$ denotes the unit step function. At position (x, y, z) the IR difference $h_{s_i r_j, d}(t; x, y, z)$ is fit to the model in (2.3.5) to yield parameters

$$h_{s_i r_j, f}(x, y, z) = [\eta(x, y, z), \tau(x, y, z), t_1(x, y, z)].$$

The database of model parameters for each grid point is denoted $\{\mathcal{H}_f(t; x_n, y_n, z_n) : n = 1, 2, \cdots, N_x \times N_y \times N_z\}$.

The procedure for fitting the IR difference to the model in (2.3.5) is to normalize $h_{s_i r_j, d}(t; x, y, z)$ to have a peak of 1. The delay $\hat{t}_1$ is estimated by detecting time instant of the peak. The signal after $\hat{t}_1$ is divided into N windows of fixed length which may overlap. The N windows of length L each are stored in columns of a matrix $\mathbf{Y}$ and the auto-covariance matrix $\mathbf{R}_{yy}$ is computed. The decay rate $\hat{\tau}$ is then estimated using MUSIC [105]. Given $\hat{t}_1$ and $\hat{\tau}$, the amplitude $\hat{\eta}$ is then estimated

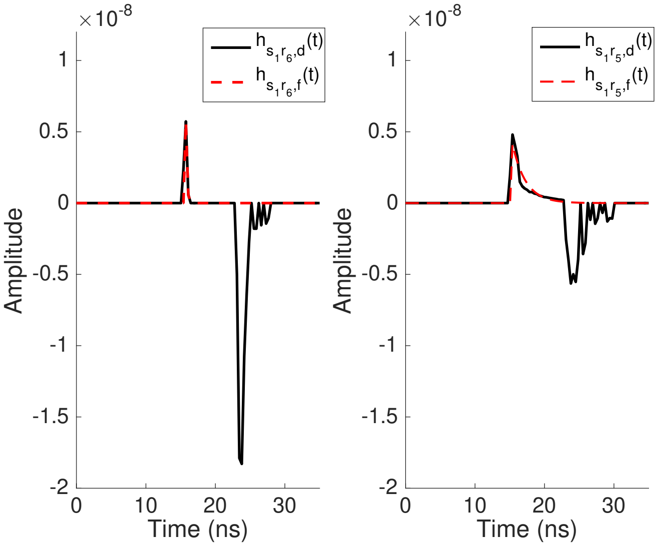


Figure 2.3: IR difference, $h_{s_i r_j,d}(t)$, and exponential model fit, $h_{s_i r_j,f}(t)$, from (left) s_1 to r_6 and (right) s_1 to r_5 by considering the scenario depicted in Figure 2.2 .

using least-squares method.

Figure 2.3 shows an example of the IR difference and the one obtained after parameter fitting. Notice the IR difference $h_{s_i r_j,d}(t)$ can be negative since it is the difference of two optical intensity IRs. The fit is tight only for the positive portion of $h_{s_i r_j,d}(t)$ which accounts for the reflections from the object while the negative components are due to shadowing. Though the fit is approximate and is often loose, the reduction in storage requirements is compelling and localization performance is quantified in Section 2.5.

2.4 Localization Algorithm

Assume that the object is located at position (x, y, z) in the room. The passive indoor localization algorithm begins with a noisy measurement of the IR difference between all source-receiver pairs with an object inside the room. It is assumed that the collection of IRs in empty room, $\mathcal{H}_e$, has already been measured in a calibration stage and is available. An estimate of the received reflections from solely the object is

$$\tilde{h}_{s_i r_j,d}\left(t;x,y,z\right) = h_{s_i r_j,o}\left(t;x,y,z\right) + n_{s_i r_j}(t) - h_{s_i r_j,e}(t) \tag{2.4.1}$$

where $\tilde{h}_{s_i r_j,d}\left(t;x,y,z\right)$ denotes the noisy IR difference and $n_{s_i r_j}\left(t\right)$ is receiver noise. The location of the object is estimated by selecting the (x_n, y_n, z_n) in $\{\mathcal{H}_d\left(t;x_n,y_n,z_n\right)\}$ which minimizes the mean square error between $\tilde{h}_{s_i r_j,d}\left(t;x,y,z\right)$ and the database. More precisely, for N_s sources and N_r receivers, the error matrix at n^{th} grid point is calculated as

$$\mathbf{E}_n = \begin{bmatrix} e_{11}^{(n)} & \cdots & e_{1N_r}^{(n)} \\ \vdots & \ddots & \vdots \\ e_{N_s 1}^{(n)} & \cdots & e_{N_s N_r}^{(n)} \end{bmatrix} \tag{2.4.2}$$

where $e_{ij}^{(n)}$ is defined as root mean square error between $\tilde{h}_{s_i r_j,d}\left(t;x,y,z\right)$ and $h_{s_i r_j,d}\left(t;x_n,y_n,z_n\right)$ i.e.

$$e_{ij}^{(n)} = \sqrt{\int\limits_0^{t_{\max}} \left|\tilde{h}_{s_i r_j,d}\left(t;x,y,z\right) - h_{s_i r_j,d}\left(t;x_n,y_n,z_n\right)\right|^2 dt} \tag{2.4.3}$$

where $t_{\max}$ denotes the maximum time used in IR acquisition.

To estimate the location, the Frobenius norm of $\mathbf{E}_n$ at each grid point is computed. The grid point at which $\|E_n\|_F$ is minimum is reported as the estimated position of the object and is denoted as $\hat{\alpha} = [\hat{x}_n, \hat{y}_n, \hat{z}_n]$. If $\alpha_o = [x, y, z]$ is the actual location of the object, the RMS error is calculated as

$$e_{\text{rms}} = \sqrt{\left|\hat{\alpha} - \alpha_o\right|^2} \tag{2.4.4}$$

An identical procedure is used for the database of IR parameters $\{\mathcal{H}_f\,(t; x_n, y_n, z_n)\}$, where e_{ij} in (2.4.3) is computed by sampling (2.3.5) with parameters $\eta(x_n, y_n, z_n)$, $\tau(x_n, y_n, z_n)$, $t_1(x_n, y_n, z_n)$ at the same rate as the measured $\tilde{h}_{s_i r_j,d}(t)$.

The accuracy of the proposed technique depends on density of the luminaires as well as the grid points. Thus, increasing the density of these increases the size of database. So there is a trade off between the accuracy and workload requirement for deployment. However, crowdsourcing of IR readings can reduce the workload.

2.5 Simulation Results

2.5.1 Simulation Environment

The room has dimensions $5\,\mathrm{m} \times 5\,\mathrm{m} \times 3\,\mathrm{m}$ and is illustrated in Figure 2.4. The total number of sources and receivers is $N_s = N_r = 9$ which are co-located on the ceiling and equally spaced. Table 2.1 presents a list of parameters used in the simulations. To simplify the simulations, positioning is simulated in two-dimensions and $z = z_h =$ 1.5 m which is the height of the object.

The sources are considered Lambertian with radiation pattern [106]

$$R\,(\phi) = \frac{m+1}{2\pi} P_s \cos^m(\phi) \tag{2.5.1}$$

where m is the Lambertian index, P_s is the power emitted by the light source and ϕ is the angle formed between the light ray and normal of the light source. The ceiling-mounted receivers are modelled as having an effective area A_r with a field-of-view (FOV) ψ_r. Moreover, the reflections of light in the room are considered diffuse.

Table 2.1: Values for Simulation Parameters

Room	Size, $5\,\mathrm{m} \times 5\,\mathrm{m} \times 3\,\mathrm{m}$ Discretization step, 10 cm Patch size, $10\,\mathrm{cm} \times 10\,\mathrm{cm}$ No. of sources and receivers, $N_s = 9$, $N_r = 9$ Reflection coefficient, $\rho_{\text{floor}} = 0.7$, $\rho_{\text{walls}} = 0.85$ Sampling rate of IR, $\Delta t = 0.35$ ns
Light Source	Transmit power, $P_s = 1$ W Lambertian index, $m = 1$
Receiver	FOV (half-angle), $\psi_r = 45\,\mathrm{deg}$ Surface area, 1 cm^2 Responsivity (γ) = 1 A/W
Object	Size, $0.3\,\mathrm{m} \times 0.3\,\mathrm{m} \times 1.5\,\mathrm{m}$ Patch size, $3\,\mathrm{cm} \times 3\,\mathrm{cm}$ Reflection coefficient, $\rho_{\text{top}} = 0.5$, $\rho_{\text{sides}} = 0.6$

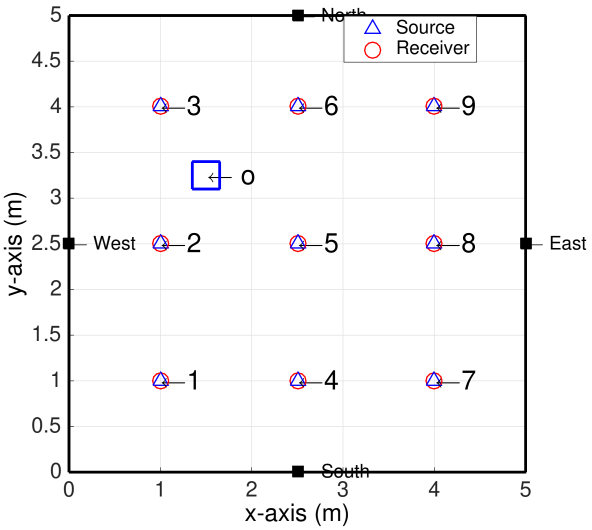


Figure 2.4: Top view of the room used in analysis. An example placement of the object is also shown, which is at $(1.5\,\text{m}, 3.25\,\text{m}, 1.5\,\text{m})$. The sources and receivers are co-located on the ceiling.

In order to simulate the IRs, the well known algorithm in [106] is used where the room is segmented in a series of non-overlapping patches (of size 10×10 cm) where the light is emitted from source s_i and received at the receiver r_j after l reflections from the floor or walls. To simplify the analysis, all IRs are simulated using a single bounce model (i.e., $l = 1$) and are sampled on a time scale of 0.35 ns. In the following the superscript '(1)' is used to denote the IRs generated in this fashion.

The signal-to-noise ratio is defined here as

$$\text{SNR} = \frac{\gamma^2 P_s^2}{\sigma_n^2} \tag{2.5.2}$$

where γ (A/W) is receiver responsivity, P_s is the power emitted by the source, and the variance of the noise at each receiver is σ_n^2. For all simulations, $\sigma_n^2 = 10^{-16}\text{A}^2$.

2.5.2 IRs with and without Object

Figure 2.5 presents an example of the IRs computed in the room with and without the presence of an object from source s_1 to receiver r_6 shown in Figure 2.2. Notice that the response of the empty room, $h^{(1)}_{s_1r_6,e}\left(t\right)$ and when the object is in the room, $h^{(1)}_{s_1r_6,o}\left(t;x,y,z_h\right)$ have a significant similarity. The IR difference between these signals, $h^{(1)}_{s_1r_6,d}\left(t;x,y,z_h\right)$, is a short time signal with a positive pulse followed by a later negative pulse. Recall that the IR difference signal may assume negative values when $h_{s_ir_j,o}(t)$ is smaller than $h_{s_ir_j,e}(t)$ at any time instant due to shadowing.

Notice that background response of the empty room is largely removed in $h^{(1)}_{s_1r_6,d}\left(t;x,y,z_h\right)$ and the early non-negative portion corresponds to the IR from the object itself. The negative pulse in $h^{(1)}_{s_1r_6,d}\left(t;x,y,z_h\right)$ is due to object shadowing and aids localization in the case of the full database $\{\mathcal{H}_d\left(t;x_n,y_n,z_n\right)\}$. In the case of the fitted database $\{\mathcal{H}_f\left(x_n,y_n,z_n\right)\}$, however, this negative pulse is clipped to aid fitting (as described in Sec. 2.3.4) and does not contribute to localization.

2.5.3 Performance Evaluation

Figures 2.6 and 2.7 show the localization performance using both a complete database $\{\mathcal{H}_d\left(x_n,y_n,z_n\right)\}$ as well as the fit database $\{\mathcal{H}_f\left(x_n,y_n,z_n\right)\}$ averaged over 1000 random positions inside the target room. For each choice of database the simulations are done for different grid sizes which correspond to different database sizes. Since the primary function of the luminaires is to provide illumination to the room, localization performance is also plotted versus source brightness variation which can arise due to dimming of the sources.

As expected, the performance degrades by decreasing the number of grid points for

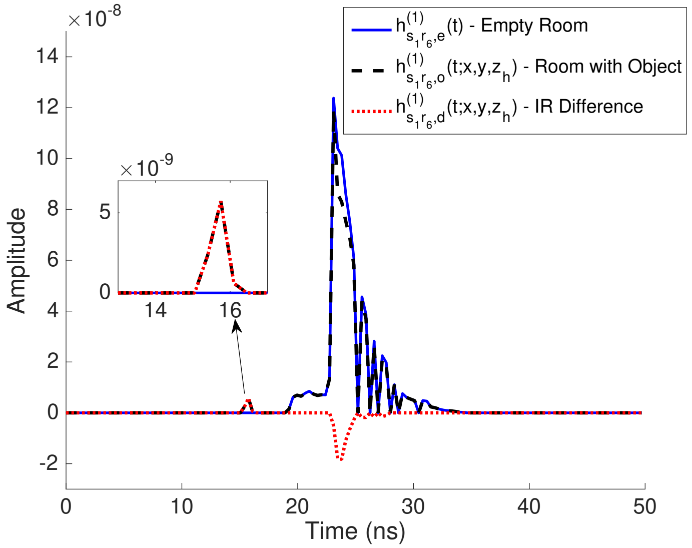


Figure 2.5: IR with and without an object inside the room along with the IR difference. The source s_1 is located at $(1\,\mathrm{m}, 1\,\mathrm{m}, 3\,\mathrm{m})$ and receiver r_6 is located at $(2.5\,\mathrm{m}, 4\,\mathrm{m}, 3\,\mathrm{m})$. The object is located at $(1.5\,\mathrm{m}, 3.25\,\mathrm{m}, 1.5\,\mathrm{m})$. The values used in simulation are shown in Table 2.1.

both techniques. However, the improvement in increasing the database sizes quickly saturates. The performance using the larger $\{\mathcal{H}_d\,(x_n, y_n, z_n)\}$ is superior to the less complex $\{\mathcal{H}_f\,(x_n, y_n, z_n)\}$. Additionally, the performance with $\{\mathcal{H}_f\,(x_n, y_n, z_n)\}$ is more sensitive to dimming and degrades more quickly as P_s is reduced. The fitted database, though requiring less storage, shows degraded performance since the extracted parameters from IR differences are sensitive to noise.

The previous results were obtained using a single IR reading for each source-receiver pair. Performance can be further improved by collecting several IR readings at a single position and averaging them to reduce the effect of noise at the expense of small added latency. Table 2.2 shows the mean and standard deviation (STD) of the RMS error for a single and an average of IR readings collected for each source-receiver pair at full brightness using complete and fitted databases.

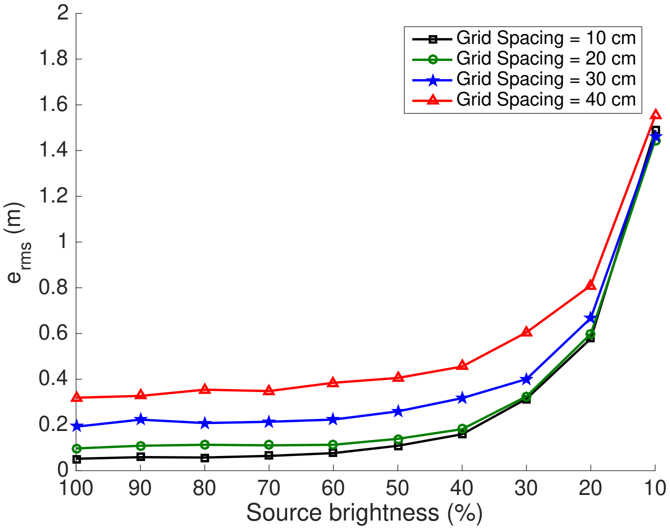


Figure 2.6: Localization error against the decreasing brightness of light sources using the complete database of IR differences $\{\mathcal{H}_d\left(t; x_n, y_n, z_n\right)\}$.

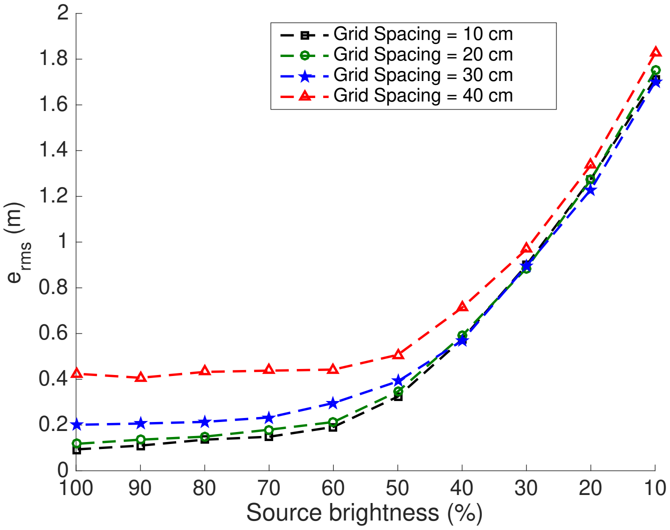


Figure 2.7: Localization error against the decreasing brightness of light sources using the fitted database of IR differences $\{\mathcal{H}_f\left(x_n, y_n, z_n\right)\}$.

Table 2.2: Mean and STD of RMS error by using complete and fitted database (grid spacing of 10 cm)

Database	**No. of collected IR readings = 1**		**No. of collected IR readings = 1000**	
	mean (cm)	**STD (cm)**	**mean (cm)**	**STD (cm)**
$\{\mathcal{H}_d\left(t; x_n, y_n, z_h\right)\}$	5.38	4.45	4.45	2.51
$\{\mathcal{H}_f\left(x_n, y_n, z_h\right)\}$	10.43	15.65	6.35	4.96

The average localization performance is particularly improved for the case of $\{\mathcal{H}_f\left(x_n, y_n, z_n\right)\}$ given the sensitivity of the fit to the noisy data. Additionally, for both approaches the variance in the location estimate is greatly reduced by averaging over 1000 IR measurements.

2.5.4 Sensitivity of $\mathcal{H}_d^{(1)}\left(t; x, y, z\right)$ to Background Variation

The sensitivity of the IR difference $\mathcal{H}_d^{(1)}\left(t; x, y, z\right)$ to a variation in the background is tested by adding a fixed element to the room. Consider a single source-receiver pair, as shown in Figure 2.8, which is heavily shadowed by the fixed element.

Denote $h_{sr,e'}^{(1)}(t)$ as the IR when fixed element is present and $h_{sr,o'}^{(1)}(t; x, y, z_h)$ when both the actual object and fixed element are present together. The IR difference in the case of a fixed element is denoted $h_{sr,d'}^{(1)}(t; x, y, z_h)$. Figure 2.8 presents the normalized mean square error between IR differences computed as

$$10 \log \left(\frac{\int\limits_0^{t_{\max}} \left(h_{sr,d}^{(1)}\left(t; x, y, z_h\right) - h_{sr,d'}^{(1)}\left(t; x, y, z_h\right) \right)^2 dt}{\int\limits_0^{t_{\max}} \left(h_{sr,d}^{(1)}\left(t; x, y, z_h\right) \right)^2 dt} \right) \qquad (2.5.3)$$

for different object locations in the room. The figure demonstrates that even in the

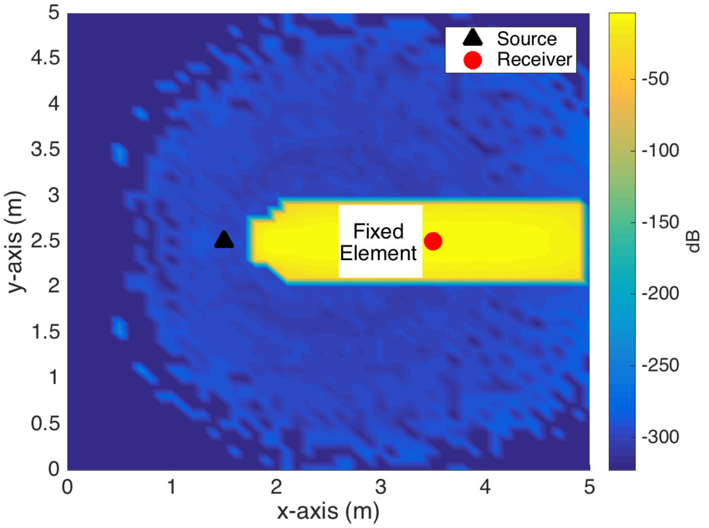


Figure 2.8: Normalized mean square error (2.5.3) between $h^{(1)}_{sr,d}(t;x,y,z_h)$ and $h^{(1)}_{sr,d'}(t;x,y,z_h)$. The fixed element has dimensions $(0.5\,\text{m}, 0.5\,\text{m}, 1.5\,\text{m})$ and is located at $(3\,\text{m}, 2.5\,\text{m}, 1.5\,\text{m})$. The source s and receiver r are located at $(1.5\,\text{m}, 2.5\,\text{m}, 3\,\text{m})$ and $(3.5\,\text{m}, 2.5\,\text{m}, 3\,\text{m})$ respectively.

presence of a major changes in the background environment, that the IR difference is essentially unchanged for a large portion of the room. The largest differences occur due to the shadowing of the fixed element. Thus, when changes to the background occur, the database needs to be updated for only those locations which fall in the shadowed areas.

2.6 Conclusions

The paper presents a new paradigm for indoor localization which does not require user intervention to localize. This passive positioning provides estimates of the user location by measuring the changes in the IR of the room between a series of LED sources and receivers on the ceiling. Although this passive approach requires a network of connected luminaires and detectors, this vision is already in the progress of being realized. Recent developments in IP connected lighting infrastructure connected via power-over-Ethernet suggest that such backbone networks will be imminently available in new installations [56]. The workload required for building the database can be significantly reduced by collecting crowdsourced readings. Our initial simulations indicate that at high brightness an object can be localized to an accuracy on the order of 5 cm which degrades as the luminaires are dimmed. Our ongoing work focuses on reducing the storage overhead of this approach as well as experimental prototypes for passive positioning.

Acknowledgment

The authors would like to thank Dr. James Reilly for his help regarding the fitting of IR parameters.

Chapter 3

Performance Bounds on Passive Indoor Positioning using Visible Light

The research work presented in this chapter appeared in our following journal article [88].

- **K. Majeed** and S. Hranilovic, "Performance Bounds on Passive Indoor Positioning Using Visible Light," in *Journal of Lightwave Technology*, vol. 38, no. 8, pp. 2190-2200, 15 April 15, 2020, doi: 10.1109/JLT.2020.2966365.

Minor modifications are done in the manuscript shown in this chapter in order to follow the thesis format i.e., formatting of equations, citations, figures, etc, however, the main content in the reproduced article is the same. The paper also has some supplementary material, attached to the journal article, which contains the details of the derivations of the derivatives required to compute the Cramér-Rao lower bound (CRLB). The contents of the supplementary material are shown in Appendix A of the thesis.

The localization object (LO) in this chapter is assumed present inside the room in order to calculate the CRLB. Furthermore, no ambient light sources are considered. This results in detection probability of 1 with 0 false alarms.

This chapter employs the same single-bounce model used in Chapter 2. However, in this chapter, theoretical performance is characterized on the positioning error of a passive indoor localization system. The room and system model considered to obtain impulse response (IR) measurements between the source-receiver pairs follow the same methodology described in Chapter 2. Furthermore, only first-order reflections are used to represent IRs between different source-receiver pairs. The IR difference is obtained by subtracting the IR without the LO in the room from the IR when the LO is present inside the room. The IR differences are zero-clipped in order to remove the negative part and exponential integrating-sphere model [92] is used to fit the zero-clipped IR differences. The amplitude, time delay and decay rate of the fitted IR model is represented as a function of the Cartesian coordinates of the LO.

The fitted model is then used to derive CRLB on the positioning error using log-likelihood function. The log-likelihood function is a function of the receiver noise and parameters of the fitted model that are functions of the position of LO. The derivatives of log-likelihood function are computed in order to obtain the Fisher information matrix (FIM) that is ultimately used to compute the analytical expression of CRLB. The tightness of the bound is evaluated by using maximum likelihood (ML) estimator that estimates position of the LO by maximizing the log-likelihood function.

This is important to note that the CRLB derived in this chapter is given as the bound on any unbiased estimator. However, the mathematical analysis is not performed in this chapter in order to verify that the ML estimator is unbiased.

The location of sources and receivers are considered fixed in the room in order to derive the bound and estimate position. The luminaires are typically fixed in the indoor area (i.e., on the ceiling) and are not moved over extended periods of time. The estimation algorithms are usually sensitive to the position of sources and receivers. Though mathematical analysis of the sensitivity due to changing location of sources and receivers is not performed, however, small changes e.g. mm or even few cm in the location do not impact the performance considerably. This is due to the fact that peaks in the IRs because of the LO shift on the order of a fraction of a nanosecond with a few cm offset in location of the sources/receivers.

3.1 Abstract

In this paper, a novel method for passive indoor localization using LED luminaires is proposed where explicit user participation is not required. This approach measures changes in the impulse response between sources and receivers and estimates a location based on optical channel sounding data. An exponential integrating-sphere model is used to represent object impulse response (OIR) from each luminaire source-receiver pair, which is obtained by subtracting impulse response (IR) of the room background (i.e., without an object) from IR when the object is present inside the room. This model is represented as a function of 3D position of the object and depends on both source and receiver parameters and the physical geometry of the room. An analytical expression of Cramér-Rao lower bound (CRLB) on the proposed passive indoor localization method is derived. The position is also estimated by using a maximum likelihood (ML) estimator which gives the position estimate by maximizing the log-likelihood function of the received noisy OIR waveforms. The results show that the signal-to-noise ratio (SNR) and number of source-receiver pairs used in the estimation play a crucial role in performance. Typical localization root-mean squared error is less than 10 cm over a broad range of light intensities and object locations.

3.2 Introduction

Indoor illumination is dominated by the use of solid-state light-emitting diode (LED) technology due to their energy-efficiency, cost-effectiveness and long life-span. In addition to their primary role as illuminators, the ubiquity of LED luminaires make them suitable to support secondary functions such as indoor communications and positioning. In particular, visible light positioning (VLP) has benefits of leveraging existing illumination installations while providing greater privacy, security and accuracy over comparable radio frequency (RF) approaches. Since light is easily contained by opaque boundaries, VLP emissions are contained to the room of interest which stands in contrast to RF-based positioning where the RF signals travel through walls and are susceptible to eavesdropping. Recent VLP-based techniques [107, 101, 108, 65, 51, 99, 37, 109, 60, 110, 111, 112] can be broadly categorized into the following categories: proximity-based, received signal strength (RSS)-based and triangulation-based approaches.

In *proximity*-based approaches [39, 100, 113, 114, 37], the LED light sources are considered to emit a unique signature which is detected by a receiver at the user location. In proximity-based techniques the luminaires are treated as beacons and receivers declare their position based on the closest beacon detected. These techniques typically have low accuracy as compared to other methods [51]. For *RSS*-based techniques [115, 116, 117, 118, 119, 59] the RSS of received optical signals are used for localization. The average optical power measured at the receiver from several sources is used to estimate receiver location by solving a set of equations which are based on indoor optical wireless channel models (e.g., [45]). A subclass of these methods (e.g., [59, 65, 120]) create a fingerprint map of RSS measurements throughout an indoor space which is constructed during an offline stage. Noisy measurements collected during the online stage are then compared to the stored fingerprint map and the grid point which gives minimum mean squared error between the noisy measurement and stored value in the fingerprint map is returned as the position estimate. These approaches require significant measurements and rely heavily on the accuracy of the underlying models of RSS measurements. *Triangulation*-based techniques [121, 111, 70, 112, 69] use time-of-arrival (TOA), time-difference-of-arrival (TDOA), or angle-of-arrival (AOA) information from the received optical signals for localization. The time-related information is extracted from the received signals from several sources in TOA and TDOA-based methods, which is then used to solve a set of equations to obtain the position estimate. The AOA-based methods typically employ complex multiple photo-detector (PD) arrays at the receiver to estimate the elevation or azimuth angle of the receiver. Some variants of PD array arrangement have also been studied, for example in [122] a circular PD array for location estimation. The major challenge of these methods is that perfect time synchronization is required between the source and receiver.

In addition to positioning techniques, there has been considerably less work to calculate performance bounds of VLP techniques [69, 64, 67, 123, 124, 125, 126, 127]. It is common to calculate the Cramér-Rao lower bound (CRLB), which is given as a lower bound on the root mean squared error (RMSE) of an unbiased estimate of unknown variables in mean squared error sense [128]. The CRLB of RSS-based techniques was calculated in [64, 61]. The authors in [67, 123] calculate CRLB for a hybrid RSS/TOA-based technique using multiple PDs at the receiver. The methods in [126, 69] calculate CRLB on TOA-based positioning techniques. All the methods mentioned earlier fall under an *active localization* paradigm, i.e., where the user estimates their position by employing a receiver with a direct line-of-sight (LOS) to the light sources.

This paper presents bounds on the performance of a fundamentally new paradigm in VLP termed *passive localization* [120]. In this approach, light sources and receivers are affixed on the ceiling (potentially co-located) and are controlled by an underlying network (e.g., power-over-ethernet or power line communication channels). Measurements of the impulse responses (IRs) between source-receiver pairs are used to estimate the location of an object in the room. In spite of the challenge of synchronizing amongst sources and receivers, the underlying network can be calibrated to provide a common time reference for measuring IRs as was done in earlier TOA and TDOA work [69, 124, 123, 129]. The measurement of IRs is done such that one source is on at a time and this process is repeated for all the remaining sources. The IR measurements for a given source-receiver pair can be done on the order of 10's of ns and can thus easily be made imperceptible to the human eye. Assuming a single object to be detected in the room, an exponential integrating-sphere model [104, 92] is used to represent the IR of the object to be localized where the parameters of the model depend on the location of the object and the geometry of the room. Using this model, an analytical expression for CRLB is derived as a lower bound on the position estimate error. The position of the object in the room is then estimated using a maximum likelihood (ML) estimator to quantify the performance in a simulation environment. This passive VLP is similar in approach to channel sounding employed in underwater acoustic channels, where measured reflected signals are used to infer position [103] as well as related work in MIMO radar with widely-separated antennas [130]. A related approach has been proposed in parallel in [76], however, they rely only on power measurements for outdoor localization scenarios.

In contrast to existing active VLP approaches, passive VLP does not require fixed beacons, specialized PD arrangements or user involvement in the localization scheme. This passive approach is especially attractive in applications where user involvement in the positioning process is not possible. As an example, consider the need to monitor a patient at home or in a hospital in order to detect an accidental fall while preserving their privacy.

The notation used in this paper are the following: scalar by x, constant by X, vector by $\mathbf{x}$, i^{th} element of a vector $\mathbf{x}$ by x_i, matrix by $\mathbf{X}$, i^{th} column of a matrix $\mathbf{X}$ by $\mathbf{x}_i$, ij^{th} entry of a matrix $\mathbf{X}$ by x_{ij}, set by $\mathcal{X}$, set cardinality by $|\mathcal{X}|$, and Euclidean norm of a vector $\mathbf{x}$ by $\|\mathbf{x}\|_2$. Vectors are considered column vectors throughout the paper and $[\cdot]^T$ denotes the transpose operation.

The remainder of the paper is organized as follows. Section 3.3 describes the proposed passive localization scenario and the model used to represent IR differences. The analytical expression for CRLB is derived in Section 3.4. The ML position estimation is obtained in Section 3.5. Numerical results are described and discussed in Section 3.6. Finally, the paper is concluded in Section 3.7.

3.3 System Model

3.3.1 Problem Formulation

Consider an object o inside the room as shown in Figure 3.1 with N_s light sources and N_r receivers affixed to the ceiling. Though shown to be co-located in Figure 3.1, which may be the least costly in terms of deployment, sources and receivers can be arbitrarily distributed on the ceiling and are assumed to be parallel to the floor.

The light sources are considered Lambertian emitters [106, 45] with radiation pattern in Watts/steradian

$$R(\phi) = \frac{m+1}{2\pi}(\beta P_s)\cos^m(\phi) \tag{3.3.1}$$

where m denotes the Lambertian index of the source, P_s denotes the total power transmitted by the source in Watts, β is a factor to control the brightness of the light source, and ϕ denotes the angle formed between the normal to the source and a distance vector from source to the object or surface (see Figure 3.1). Receivers are assumed to have field-of-view (FOV) ψ and effective area A_r. Moreover, surfaces of the room and object are modelled as diffuse reflectors.

Consider the example of light rays emerging from source s_i and diffusely reflecting from an object, the floor and walls as shown in Figure 3.1. The positions of source s_i and receiver r_j are represented by vectors $\mathbf{x}_i^{(s)} = [x_{s_i}\ y_{s_i}\ z_{s_i}]^T$ and $\mathbf{x}_j^{(r)} = \left[x_{r_j}\ y_{r_j}\ z_{r_j}\right]^T$ respectively. The object is modelled as a cuboid with a square top denoted by $\mathcal{S}_o$ where $\mathbf{x}_o = [x_o\ y_o\ z_o]^T$ represents the unknown position of the centre of $\mathcal{S}_o$.

Define the *object impulse response* (OIR), $h_{s_i r_j}(t, \mathbf{x}_o)$, as the IR between s_i and r_j considering reflected rays solely from $\mathcal{S}_o$ located at $\mathbf{x}_o$. Using the approach in [106], the OIR can be written as

$$\begin{aligned} h_{s_i r_j}(t, \mathbf{x}_o) = \sum_{o_k \in \mathcal{S}_o} (\beta P_s) \frac{(m+1)\rho_o A_{r_j} \Delta A_{o_k}}{2\pi d^2_{s_i o_k} d^2_{o_k r_j}} \cos^m \phi_{s_i o_k} \\ \cos\theta_{o_k s_i} \cos\phi_{o_k r_j} \cos\theta_{r_j o_k} \delta\left(t - \frac{d_{s_i o_k} + d_{o_k r_j}}{c}\right) \end{aligned} \tag{3.3.2}$$

where o_k is the k^{th} patch of $\mathcal{S}_o$ so that $\bigcup_k o_k = \mathcal{S}_o$, ρ_o is reflectivity, A_{r_j} is area of the receiver r_j, ΔA_{o_k} is area of o_k, and c is the speed of light. In order to compute the

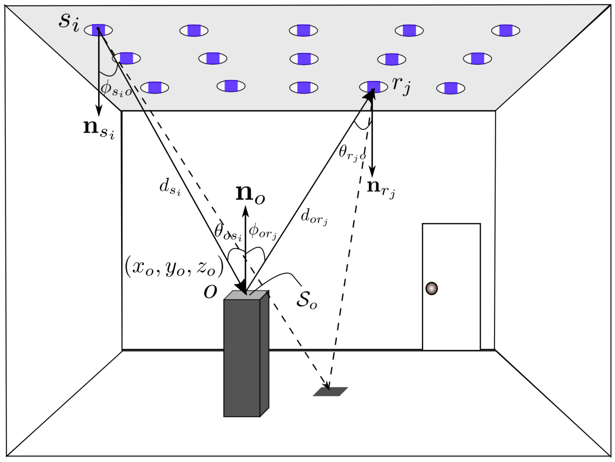


Figure 3.1: A room with light sources and receivers affixed on the ceiling. The sources and receivers are co-located. An object (i.e., a cuboid) is also present. A light ray originating from source s_i is reflected from the object (or floor and walls) and received at receiver r_j.

OIR in (3.3.2), the object surface is discretized into small patches as in [106]. There is a unique path for each ray from source s_i to patch o_k and finally to receiver r_j. The OIR in (3.3.2) is then the sum of reflected rays from all the patches on the object's top surface.

In practice, the OIR can be approximated by measurements of the IR between sources and receivers in the room. Let $h_{s_ir_j,e}(t)$ denote the IR of an empty room from s_i to r_j and similarly, $h_{s_ir_j,o}(t, \mathbf{x}_o)$ denotes the IR when the object is present inside the room. Define the *IR difference*, $\bar{h}_{s_ir_j}(t, \mathbf{x}_o)$, as

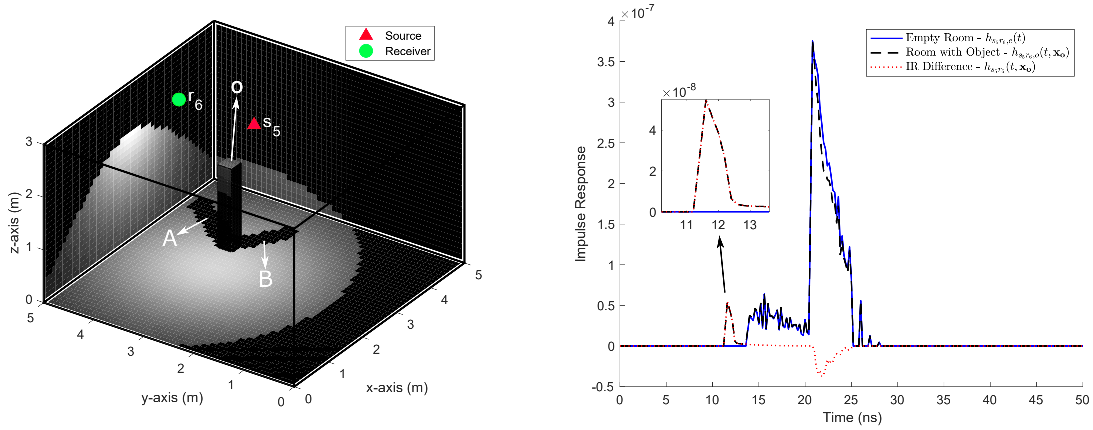
$$\begin{aligned} \bar{h}_{s_ir_j}(t, \mathbf{x}_o) &= h_{s_ir_j,o}(t, \mathbf{x}_o) - h_{s_ir_j,e}(t) \\ &\approx h_{s_ir_j}(t, \mathbf{x}_o). \end{aligned} \tag{3.3.3}$$

In practice the IR differences between source-receiver pairs are taken as an estimate of the OIRs and can be obtained independently and imperceptibly by turning each source-receiver pair on one-at-a-time under the control of an underlying coordination network. The task of the passive localization algorithm is to infer the position $\mathbf{x}_o$ from a collection of IR difference measurements of $\{\bar{h}_{s_i r_j}(t, \mathbf{x}_o)\}$ over s_i and r_j.

3.3.2 Comparison of $\bar{h}_{s_i r_j}(t, \mathbf{x}_o)$, $h_{s_i r_j, o}(t, \mathbf{x}_o)$ and $h_{s_i r_j, e}(t)$

The presence of object inside the room modifies the IR. The formation of shadows due to an object o are shown in Figure 3.2 by considering a single source-receiver pair at a location in an example room scenario. Figure 3.2a shows two shadows formed on the floor. Shadow A is formed due to obstruction of the light path due to the object top and sides. Shadow B, on the other hand, is formed due to shadowing of the receiver because these areas are not visible to the receiver. Figure 3.2b shows a comparison of IR with and without an object present inside the room, i.e., $h_{s_5 r_6, o}(t, \mathbf{x}_o)$ and $h_{s_5 r_6, e}(t)$. Notice that the presence of an object creates a small peak in IR earlier in time because the top surface of the object is at a shorter distance to the receiver as compared to other surfaces of the object or the room. The difference of two IRs is also shown by a dotted red line. The negative part in the difference corresponds to the shadowing and occurs later in time. An important point to note here is that the red curve is the difference of IRs and not an IR.

In this work only first-order reflections from all surfaces in the room and object are simulated (i.e., light travels from source to receiver through a single reflection). Higher order reflections carry very little power and are often buried in the noise floor of the receiver. Notice also that higher order reflections occur at a later time after the first-order reflections are received and can be easily removed from measured IRs [106]. Thus, for the purposes of passive localization, the IR difference $\bar{h}_{s_i r_j}(t, \mathbf{x}_o)$ is



(a) (b)

Figure 3.2: The object of fixed height ($z_o = 1.6$ m) is located at $(3.25\,\text{m}, 3.5\,\text{m}, 1.6\,\text{m})$ with all other parameters as defined in Section 3.6. Source s_5 and receiver r_6 are located at $(2.5\,\text{m}, 2.5\,\text{m}, 3\,\text{m})$ and $(2.5\,\text{m}, 4\,\text{m}, 3\,\text{m})$ respectively: (a) Shadow plot with presence of object and (b) comparison of IR with and without an object inside the room as well as IR difference (first-order reflections considered only).

dominated by the first-order reflections from the object surfaces.

3.3.3 Modeling the OIR

Consider modelling the first-order reflections from the object, i.e., the OIR in (3.3.2), with the continuous exponential model [104, 92]

$$\tilde{h}_{s_i r_j}(t, \mathbf{x}_o) = \eta_{s_i r_j}(\mathbf{x}_o)\, e^{-\frac{\left(t - t_{s_i r_j}(\mathbf{x}_o)\right)}{\tau_{s_i r_j}(\mathbf{x}_o)}}\, u_{\sigma_a,\varepsilon}\left(t - t_{s_i r_j}(\mathbf{x}_o)\right) \tag{3.3.4}$$

where the parameters $\eta_{s_i r_j}(\mathbf{x}_o)$, $t_{s_i r_j}(\mathbf{x}_o)$ and $\tau_{s_i r_j}(\mathbf{x}_o)$ are related to $\mathbf{x}_o$ in the following. The function $u_{\sigma_a,\varepsilon}(t)$ is the Gaussian cumulative distribution function

$$u_{\sigma_a,\varepsilon}(t) = \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{t+\varepsilon}{\sqrt{2\sigma_a^2}}\right)\right).$$

The values of ε and σ_a are free parameters chosen to ensure a good fit to the OIR in practice as described in Section 3.6. Notice that for $\varepsilon, \sigma_a \to 0$, that (3.3.4) approaches the well-known integrating sphere model for indoor optical intensity impulse responses [92].

The parameters in the OIR model in (3.3.4) are a function of unknown position vector $\mathbf{x}_o$, the location of source-receiver pairs, Lambertian parameters of the sources, receiver's parameters, reflectivity of the room surfaces, and the reflectivity of the object surfaces. The distances and angles shown in Figure 3.1 are defined as

$$\begin{aligned} d_{s_i o} &= \left\| \mathbf{x}_o - \mathbf{x}_i^{(s)} \right\|_2 \\ d_{o r_j} &= \left\| \mathbf{x}_j^{(r)} - \mathbf{x}_o \right\|_2 \\ \phi_{s_i o} = \theta_{o s_i} &= \cos^{-1}\left(\frac{z_{s_i} - z_o}{\left\| \mathbf{x}_o - \mathbf{x}_i^{(s)} \right\|_2} \right) \\ \phi_{o r_j} = \theta_{r_j o} &= \cos^{-1}\left(\frac{z_{r_j} - z_o}{\left\| \mathbf{x}_j^{(r)} - \mathbf{x}_o \right\|_2} \right) \end{aligned} \tag{3.3.5}$$

where the angles $\phi_{s_i o}$, $\theta_{o s_i}$, $\phi_{o r_j}$, $\theta_{r_j o}$ are defined with respect to normals $\mathbf{n}_{s_i}$, $\mathbf{n}_o$ and $\mathbf{n}_{r_j}$.

The amplitude parameter $\eta_{s_i r_j}(\mathbf{x}_o)$ in (3.3.4) has the same form as the scaling factor in (3.3.2) and is related to $\mathbf{x}_o$ as

$$\eta_{s_i r_j}(\mathbf{x}_o) = (\beta P_s) \frac{(m+1)\, \rho_o A_{r_j} A_{o,\text{eff}}}{2\pi} \times \frac{(z_{s_i} - z_o)^{m+1} (z_{r_j} - z_o)^2}{\left\| \mathbf{x}_o - \mathbf{x}_i^{(s)} \right\|_2^{m+3} \left\| \mathbf{x}_j^{(r)} - \mathbf{x}_o \right\|_2^4}. \tag{3.3.6}$$

The value of $A_{o,\text{eff}}$ is tuned to be a fixed proportion of the area of $\mathcal{S}_o$ in order to have a good fit of $\tilde{h}_{s_i r_j}(t, \mathbf{x}_o)$ to $h_{s_i r_j}(t, \mathbf{x}_o)$ as described in Section 3.6.

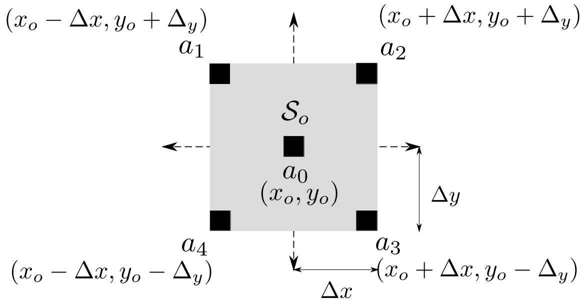


Figure 3.3: Top view of the object $\mathcal{S}_o$ showing x-y coordinates of its center a_0 and corners a_1 to a_4. The dimensions $2\Delta x \times 2\Delta y$.

The time delay $t_{s_i r_j}(\mathbf{x}_o)$ with respect to the center of $\mathcal{S}_o$ is defined as

$$t_{s_i r_j}(\mathbf{x}_o) = \frac{\left\|\mathbf{x}_o - \mathbf{x}_i^{(s)}\right\|_2}{c} + \frac{\left\|\mathbf{x}_j^{(r)} - \mathbf{x}_o\right\|_2}{c}. \tag{3.3.7}$$

This is the total time-of-flight for the signal to travel from s_i to r_j after a first reflection from the object.

The delay spread $\tau_{s_i r_j}(\mathbf{x}_o)$ is a function of $\mathcal{S}_o$ and depends on its orientation with respect to s_i and r_j. Consider the top view of $\mathcal{S}_o$ in Figure 3.3 where the dimensions of $\mathcal{S}_o$ are $2\Delta x \times 2\Delta y$. In this work, the delay spread is estimated as the root mean squared (RMS) value of the delays from the four corners of $\mathcal{S}_o$ with respect to the centre at a_0. More precisely,

$$\begin{aligned}\tau_{s_i r_j}(\mathbf{x}_o) = \frac{1}{2}\times \Big(&\left(t_{s_i r_j}^{(a_1)}(\mathbf{x}_o) - t_{s_i r_j}(\mathbf{x}_o)\right)^2 + \left(t_{s_i r_j}^{(a_2)}(\mathbf{x}_o) - t_{s_i r_j}(\mathbf{x}_o)\right)^2 \\ &+\left(t_{s_i r_j}^{(a_3)}(\mathbf{x}_o) - t_{s_i r_j}(\mathbf{x}_o)\right)^2 + \left(t_{s_i r_j}^{(a_4)}(\mathbf{x}_o) - t_{s_i r_j}(\mathbf{x}_o)\right)^2\Big)^{\frac{1}{2}}\end{aligned} \tag{3.3.8}$$

where $t_{s_i r_j}(\mathbf{x}_o)$, $t_{s_i r_j}^{(a_1)}(x_o)$, $t_{s_i r_j}^{(a_2)}(x_o)$, $t_{s_i r_j}^{(a_3)}(x_o)$, and $t_{s_i r_j}^{(a_4)}(x_o)$ denote time delays from the center a_0 and corners a_1, a_2, a_3 and a_4 of $\mathcal{S}_o$ respectively, as shown in Figure 3.3.

The expression for $t_{s_i r_j}^{(a_1)}(\mathbf{x}_o)$ is given explicitly in (3.3.9) in terms of $\mathbf{x}_o$. The time delays for the other corners can be defined in the similar manner.

$$\begin{aligned} t_{s_i r_j}^{(a_1)}(\mathbf{x}_o) = & \frac{\sqrt{\left((x_o - \Delta x) - x_{s_i}\right)^2 + \left((y_o + \Delta y) - y_{s_i}\right)^2 + \left(z_o - z_{s_i}\right)^2}}{c} \\ & + \frac{\sqrt{\left(x_{r_j} - (x_o - \Delta x)\right)^2 + \left(y_{r_j} - (y_o + \Delta y)\right)^2 + \left(z_{r_j} - z_o\right)^2}}{c} \end{aligned} \quad (3.3.9)$$

3.3.4 Comparison of h, $\bar{h}$ and $\tilde{h}$

An example of the OIR, $h_{s_i r_j}(t, \mathbf{x}_o)$, along with the IR difference, $\bar{h}_{s_i r_j}(t, \mathbf{x}_o)$, and model OIR $\tilde{h}_{s_i r_j}(t, \mathbf{x}_o)$ are shown in Figure 3.4 at a location in a prototypical room (defined in Section 3.6). There is close agreement between the true OIR and the model $\tilde{h}_{s_i r_j}(t, \mathbf{x}_o)$. The estimate of the OIR as measured by the IR difference, $\bar{h}_{s_i r_j}(t, \mathbf{x}_o)$ in (3.3.3), also corresponds closely to the OIR, however, has negative amplitudes (at ≈ 20 ns) which occur after the reflections from the object (at ≈ 10 ns). These negative amplitudes arise in $\bar{h}_{s_i r_j}(t, \mathbf{x}_o)$ since portions of the room are not visible to the receiver due to shadowing of the object as discussed in Section 3.3.2 and shown in Figure 3.2. An important point to emphasize is that $\bar{h}_{s_i r_j}(t, \mathbf{x}_o)$ in (3.3.3), is not an IR but rather the difference of IRs with and without the object in the room. As described in Section 3.5, in practice the negative amplitude shadowing artifacts in $\bar{h}_{s_i r_j}(t, \mathbf{x}_o)$ are clipped at zero to provide a good estimate of the OIR.

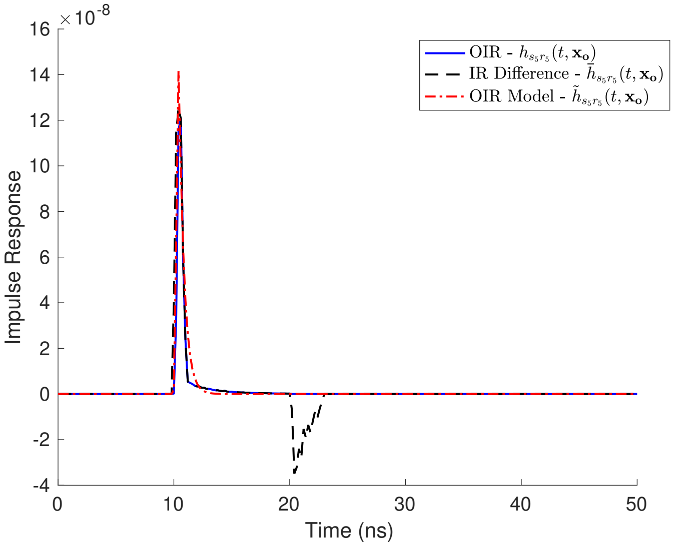


Figure 3.4: Simulated OIR (3.3.2), IR difference (3.3.3) and modelled OIR (3.3.4) from source s_5 to receiver r_5 shown in Figure 3.5. The object is located at $(3\,\text{m}, 3\,\text{m}, 1.6\,\text{m})$ with all other parameters as defined in Section 3.6.

3.4 Cramér-Rao Lower Bound

The CRLB on unbiased estimate $\hat{\mathbf{x}}_o$ of position vector $\mathbf{x}_o$ is derived in this section. The derivation of the bound assumes that the OIR recorded at each receiver has support over the time interval $t = \{0, t_{\max}\}$ typically on the order of 10's of ns. The selection of $t_{\max}$ can be related to the geometry of the room in order to ensure that first-order reflections from the object, which are used to define the OIR in (3.3.2), are included. Receivers are assumed to share a common time reference to record the responses emitted from each source. In general, consider that a subset of the measured OIRs are used to estimate position where $\mathcal{D}$ is defined as the collection of (s_i, r_j) pairs used in position estimation. The number of OIR measurements used is a metric of the complexity of the positioning and $\mathcal{D}$ can in practice be selected in many ways (e.g., using K OIRs of highest energy).

The measured OIR from source s_i to receiver r_j can be modelled as

$$z_{s_i r_j}(t) = \tilde{h}_{s_i r_j}(t, \mathbf{x}_o) + n_{s_i r_j}(t) \tag{3.4.1}$$

where $n_{s_i r_j}(t)$ is additive white Gaussian noise with zero-mean and variance σ^2 which is independent of all other source-receiver pairs. Define $\mathcal{Z} = \{z_{s_i,r_j}(t) | (s_i, r_j) \in \mathcal{D}\}$.

The log-likelihood function, considering the model in (3.4.1), using a subset of $|\mathcal{D}|$ source-receiver pairs is [128]

$$\Lambda(\mathcal{Z}, \mathbf{x}_o) = C - \sum_{(s_i,r_j)\in\mathcal{D}} \frac{1}{2\sigma^2} \int_0^{t_{\max}} \left(z_{s_i r_j}(t) - \tilde{h}_{s_i r_j}(t, \mathbf{x}_o) \right)^2 dt \tag{3.4.2}$$

where C is a constant that does not depend on $\mathbf{x}_o$. Define Fisher information matrix (FIM) $\mathbf{J}(\mathbf{x}_o)$ as

$$\mathbf{J}(\mathbf{x}_o) = E\left\{ \left(\nabla_{\mathbf{x}_o} \Lambda(\mathcal{Z}, \mathbf{x}_o)\right) \left(\nabla_{\mathbf{x}_o} \Lambda(\mathcal{Z}, \mathbf{x}_o)\right)^T \right\} \tag{3.4.3}$$

where $\nabla_{\mathbf{x}_o} \Lambda(\mathcal{Z}, \mathbf{x}_o)$ is gradient of $\Lambda(\mathcal{Z}, \mathbf{x}_o)$ with respect to $\mathbf{x}_o$ and defined as

$$\nabla_{\mathbf{x}_o} \Lambda(\mathcal{Z}, \mathbf{x}_o) = \begin{bmatrix} \frac{\partial \Lambda(\mathcal{Z}, \mathbf{x}_o)}{\partial x_o} & \frac{\partial \Lambda(\mathcal{Z}, \mathbf{x}_o)}{\partial y_o} & \frac{\partial \Lambda(\mathcal{Z}, \mathbf{x}_o)}{\partial z_o} \end{bmatrix}^T. \tag{3.4.4}$$

The CRLB on mean-squared error of the unbiased estimate $\hat{\mathbf{x}}_o$ of $\mathbf{x}_o$ is given as inverse of $\mathbf{J}(\mathbf{x}_o)$ i.e.,

$$E\left\{ (\hat{\mathbf{x}}_o - \mathbf{x}_o)(\hat{\mathbf{x}}_o - \mathbf{x}_o)^T \right\} \succeq \mathbf{J}(\mathbf{x}_o)^{-1}. \tag{3.4.5}$$

where $\mathbf{M} \succeq \mathbf{N}$ denotes that $(\mathbf{M} - \mathbf{N})$ is positive semi-definite. Given the dimensions of $\mathbf{x}_o$, in order for $\mathbf{J}(\mathbf{x}_o)$ to be invertible OIR measurements from at least three source-receiver pairs are required.

From (3.4.2) and (3.4.3) the $(k,l)^{\text{th}}$ entry of matrix $\mathbf{J}(\mathbf{x}_o)$ can be expanded as

$$J_{kl}(\mathbf{x}_o) = \sum_{(s_i,r_j)\in\mathcal{D}} \frac{1}{\sigma^2} \int_0^{t_{\max}} \frac{\partial \tilde{h}_{s_ir_j}(t,\mathbf{x}_o)}{\partial \mathbf{x}_{o,k}} \frac{\partial \tilde{h}_{s_ir_j}(t,\mathbf{x}_o)}{\partial \mathbf{x}_{o,l}} dt \quad k,l = 1,2,3 \tag{3.4.6}$$

The partial derivatives in (3.4.6) with respect to k^{th} element of $\mathbf{x}_o$ can be expanded as shown in (3.4.7).

$$\begin{aligned} \frac{\partial \tilde{h}_{s_ir_j}(t,\mathbf{x}_o)}{\partial \mathbf{x}_{o,k}} &= -\eta_{s_ir_j}(\mathbf{x}_o) \frac{\partial}{\partial \mathbf{x}_{o,k}} \left\{ t_{s_ir_j}(\mathbf{x}_o) \right\} e^{-\frac{\left(t - t_{s_ir_j}(\mathbf{x}_o)\right)}{\tau_{s_ir_j}(\mathbf{x}_o)}} \delta_{\sigma_a,\varepsilon}\left(t - t_{s_ir_j}(\mathbf{x}_o)\right) \\ &+ \frac{\partial}{\partial \mathbf{x}_{o,k}} \left\{ \eta_{s_ir_j}(\mathbf{x}_o) e^{-\frac{\left(t - t_{s_ir_j}(\mathbf{x}_o)\right)}{\tau_{s_ir_j}(\mathbf{x}_o)}} \right\} u_{\sigma_a,\varepsilon}\left(t - t_{s_ir_j}(\mathbf{x}_o)\right) \quad \text{for} \quad k = 1,2,3 \end{aligned} \tag{3.4.7}$$

Define $\delta_{\sigma_a,\varepsilon}(t) \triangleq \frac{1}{\sigma_a\sqrt{2\pi}} e^{-\frac{(t+\varepsilon)^2}{2\sigma_a^2}}$ as the derivative of $u_{\sigma_a,\varepsilon}(t)$. These partial derivatives are derived in Appendix 3.A.

The CRLB can be computed as

$$\text{CRLB} = \sqrt{\text{tr}\left\{ J(\mathbf{x}_o)^{-1} \right\}} \tag{3.4.8}$$

where $\text{tr}(\cdot)$ denotes trace of a square matrix, which is a sum of diagonal entries.

3.5 Position Estimation

In an analogous fashion to (3.4.1), define the noisy measurement of the OIR between s_i and r_j as

$$\bar{z}_{s_i r_j}(t) = \bar{h}^{+}_{s_i r_j}(t, \mathbf{x}_o) + n_{s_i r_j}(t) \tag{3.5.1}$$

and where $\bar{\mathcal{Z}}$ denotes the collection of such measurements. The signal $\bar{h}^{+}_{s_i r_j}(t, \mathbf{x}_o)$ is the measured IR difference in (3.3.3) where all negative amplitudes are clipped to zero. Define $\hat{\mathbf{x}}_o$ as the ML estimate of $\mathbf{x}_o$ obtained by maximizing the argument of the log-likelihood function $\Lambda\left(\bar{\mathcal{Z}}, \mathbf{x}_o\right)$ defined analogously to (3.4.2). Since C and measurements $\bar{\mathcal{Z}}$ do not depend on position, $\Lambda\left(\bar{\mathcal{Z}}, \mathbf{x}_o\right)$ can be simplified as

$$\Lambda\left(\bar{\mathcal{Z}}, \mathbf{x}_o\right) = \sum_{(s_i, r_j) \in \mathcal{D}} \left(\frac{1}{\sigma^2} \int_0^{t_{\max}} \bar{z}_{s_i r_j}(t)\, \tilde{h}_{s_i r_j}(t, \mathbf{x}_o)\, dt - \frac{1}{2\sigma^2} \int_0^{t_{\max}} \tilde{h}^2_{s_i r_j}(t, \mathbf{x}_o)\, dt \right) \tag{3.5.2}$$

In practice, the measured IR difference in (3.5.1) will be a bandlimited sampled noisy waveform. In the computation of $\hat{\mathbf{x}}_o$, we consider sinc interpolation to relate the samples of $\bar{z}_{s_i r_j}(t)$ to its samples.

The ML estimate $\hat{\mathbf{x}}_o$ is given as

$$\hat{\mathbf{x}}_o = \arg\max_{\mathbf{x}'_o} \sum_{(s_i, r_j) \in \mathcal{D}} \left(\frac{1}{\sigma^2} \int_0^{t_{\max}} \bar{z}_{s_i r_j}(t)\, \tilde{h}_{s_i r_j}(t, \mathbf{x}'_o)\, dt - \frac{1}{2\sigma^2} \int_0^{t_{\max}} \tilde{h}^2_{s_i r_j}(t, \mathbf{x}'_o)\, dt \right). \tag{3.5.3}$$

Notice that the CRLB developed in Section 3.4 is a lower bound on the positioning performance using the ML estimator (3.5.3).

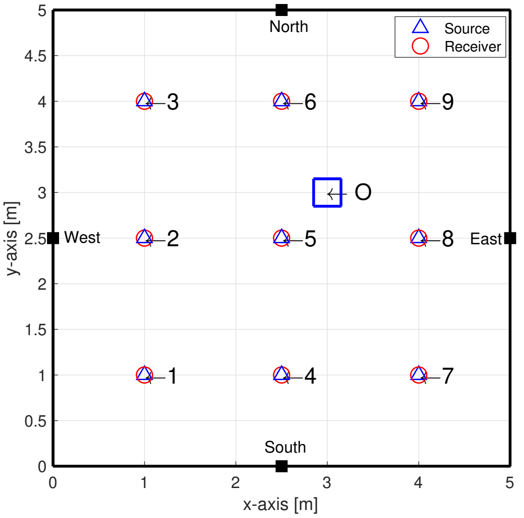


Figure 3.5: Top view of the room showing light sources and receivers affixed to the ceiling, which are co-located. The total number of sources and receivers are $N_s = 9$ and $N_r = 9$ respectively. The dimensions of the room are 5m × 5m × 3m. A cuboid object is located at $(3\,\mathrm{m}, 3\,\mathrm{m}, 1.6\,\mathrm{m})$.

3.6 Numerical Results

3.6.1 Simulation Setup

The room considered for simulations is shown in Figure 3.5 along with an example placement of an object. The values used for several simulation parameters are shown in Table 3.1.

The OIR estimate $\bar{h}_{s_i r_j}(t, \mathbf{x}_o)$ in (3.3.3) is simulated by finding $h_{s_i r_j,o}(t, \mathbf{x}_o)$ and $h_{s_i r_j,e}(t)$ using the technique in [106] considering only first-order reflections. The surfaces of room and object are divided into patches of small size, given in Table 3.1, in order to obtain $\bar{h}_{s_i r_j}(t, \mathbf{x}_o)$. Define ρ_{floor}, ρ_{wall}, ρ_{top}, and ρ_{side} as reflectivity of floor, walls, object's top surface and sides respectively. A value for $A_{o,\text{eff}}$ is found using many trial runs of OIR measurements from different source-receiver pairs and several

Table 3.1: Simulation Parameters

Room	Size, $5\,\text{m} \times 5\,\text{m} \times 3\,\text{m}$ Patch size, $10\,\text{cm} \times 10\,\text{cm}$ No. of sources and receivers, $N_s = 9$, $N_r = 9$ Reflection coefficient, $\rho_{\text{floor}} = 0.7$, $\rho_{\text{wall}} = 0.85$ Sampling rate of IR, $\Delta t = 0.2$ ns
Light Source	Transmit power, $P_s = 3$ W Lambertian index, $m = 1$
Receiver	FOV (half-angle), $\psi_r = 45\,\text{deg}$ Surface area, $A_r = 1\text{cm}^2$ Responsivity, $\gamma = 1\text{A/W}$ Noise variance, $\sigma^2 = 10^{-16}\text{A}^2$
Object	Size, $0.3\,\text{m} \times 0.3\,\text{m} \times 1.6\,\text{m}$ Patch size, $3\,\text{cm} \times 3\,\text{cm}$ Reflection coefficient, $\rho_{\text{top}} = 0.5$, $\rho_{\text{side}} = 0.6$
Other parameters	Gaussian approx. $\epsilon = 0.2$ ns and $\sigma_a^2 = 0.01\text{ns}^2$ Object top's effective area, $A_{o,\text{eff}} = 30\%$ of A_o

locations in the room to ensure a good fit with the model.

The signal-to-noise (SNR) is defined similarly to [131] as

$$\mathrm{SNR} = \frac{\gamma^2 (\beta P_s)^2}{\sigma^2} \tag{3.6.1}$$

where γ denotes the responsivity of the receiver, which is given in A/W, P_s denotes the total power emitted by a source in Watts, $\beta = \frac{P}{P_s}$ is a factor controlling the total power emitted where $\beta \in \{0.1, 1\}$ i.e., variation from 10% to 100% brightness. The receiver noise has variance σ^2 and is assumed to be the same for all receivers.

For performance evaluation, the root mean square error (RMSE) between the object's estimated and actual position is obtained for the ML estimate, which is given as

$$\mathrm{RMSE} = \sqrt{(\hat{x}_o - x_o)^2 + (\hat{y}_o - y_o)^2 + (\hat{z}_o - z_o)^2}. \tag{3.6.2}$$

The RMSE curves shown in the following subsections are obtained after averaging over several Monte Carlo (MC) runs for each point in the following results.

3.6.2 Variation of luminaire brightness

Figure 3.6 shows RMSE of the ML estimator plotted against increasing brightness of each source i.e., β is varied from 0.1 to 1. As is expected, the performance is poor at low brightness and improves by increasing brightness of the sources. The RMSE curves are also plotted in the figure for a differing number of source-receiver pairs, i.e., varying $|\mathcal{D}|$. Notice that better performance is achieved by increasing the number of pairs used at low brightness levels. However, the performance saturates quickly and shows little improvement after a certain number of pairs are used. Additionally,

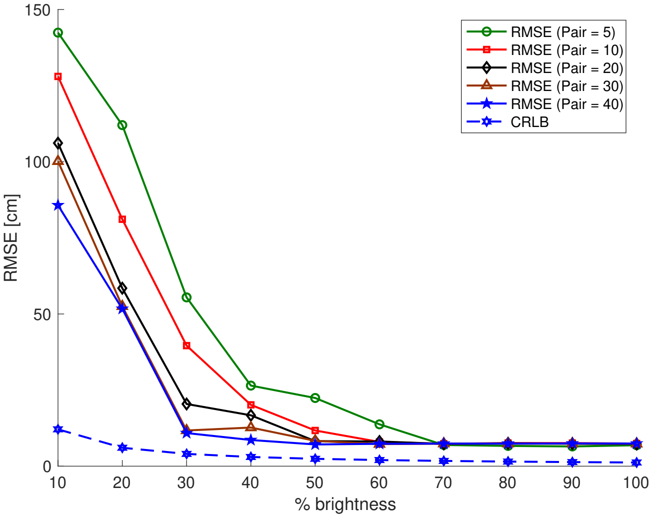
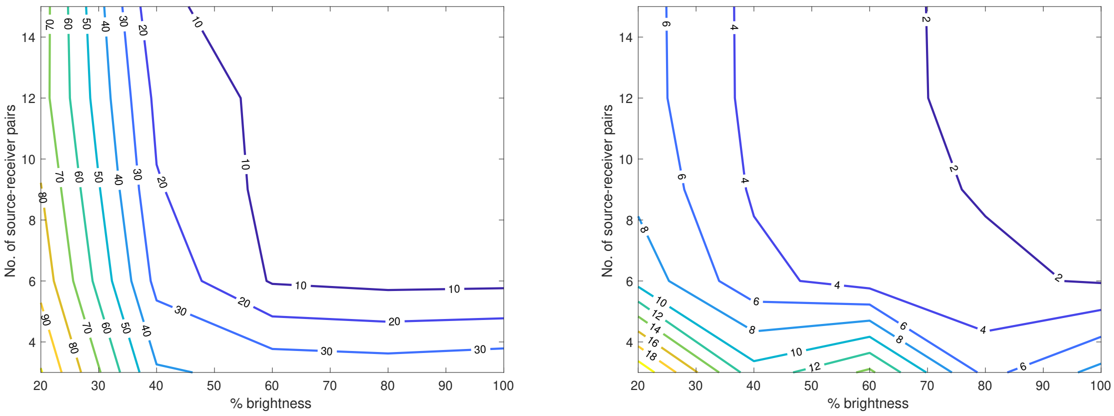


Figure 3.6: RMSE against the increasing brightness of sources by considering different number of source-receiver pairs at a time. The object is located at $(3\,\text{m}, 3\,\text{m}, 1.6\,\text{m})$ and RMSE estimated from 100 MC runs.

at high brightness, the RMSE also saturates regardless of the number of pairs used. Thus, the same level of performance can be achieved by using a smaller number of pairs at high brightness. This supports the fact that increasing the number of pairs adds up lower energy signals in building up the ML cost function, which do not aid substantially in localization.

The figure also shows a plot of the CRLB from (3.4.8) as a lower bound on the RMSE performance. The CRLB curve is obtained by using all source-receiver pairs and varying β. It should be noted that using all pairs in CRLB evaluation does not imply that all pairs contribute equally to the value of CRLB. This is because some of the receivers can not see the object since it does not fall into receiver's FOV and thus $\bar{h}_{s_i r_j}(t, \mathbf{x}_o)$ consists only of noise.



(a) (b)

Figure 3.7: Contour plots of performance against source brightness and the number of OIR measurement pairs for the case of an object located at $(3\,\mathrm{m}, 3\,\mathrm{m}, 1.6\,\mathrm{m})$: (a) RMSE [cm] (estimated from 50 MC runs.), (b) CRLB [cm].

3.6.3 Variation of number of source-receiver pairs

Figure 3.7 shows contour plots of the RMSE and CRLB against the increasing brightness of sources and number of source-receiver pairs used in estimation. From Figure 3.7a, while the RMSE is improved both by increasing the brightness of the sources and number of pairs measured, the performance is most sensitive to the choice of β. For example, at low brightness the performance is little improved even after including a considerable number of pairs. At high brightness and using a large number of measured pairs the performance, both in terms of RMSE and CRLB, becomes saturated. It is observed that the performance of the positioning algorithm is typically dominated by few high energy measured pairs which have a small time-of-flight from the corresponding sources to receivers.

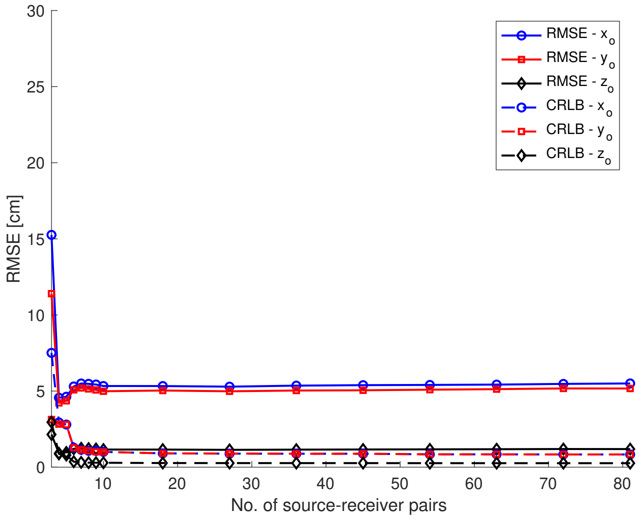
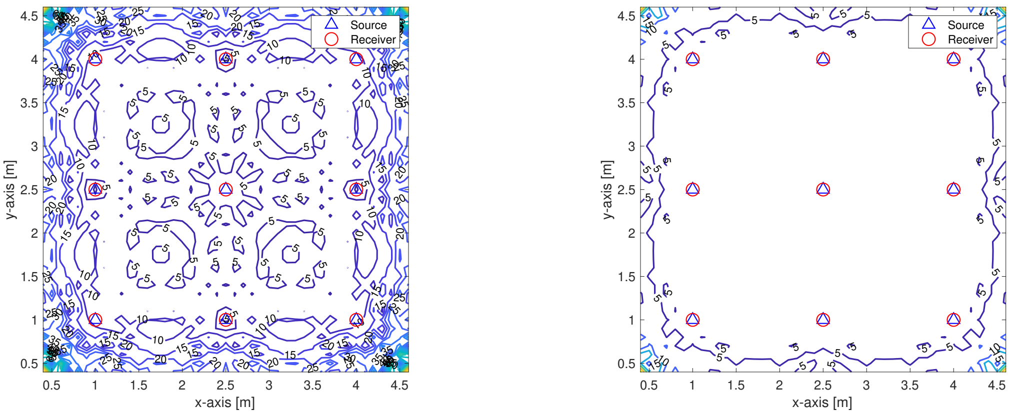


Figure 3.8: RMSE and CRLB on RMSE of individual coordinates against the increasing number of source-receiver pairs for the object located at $(3\,\mathrm{m}, 3\,\mathrm{m}, 1.6\,\mathrm{m})$. The number of pairs are varied from 3 to 81, $\beta = 1$ and RMSE results estimated from 20 MC runs.

3.6.4 CRLB on RMSE of individual coordinates

The RMSE and CRLB on individual coordinates is evaluated by increasing the number of source-receiver pairs. Figure 3.8 shows plots of RMSE and CRLB on RMSE of individual coordinates. The number of pairs are varied from 3 to all possible (81 in this case). Notice that the results in x an y axes largely mirror one another as is expected. Additionally, notice that the positioning estimation for z-coordinate is more accurate than the other coordinates. This is consistent with the fact that the Lambertian model has uniaxial symmetry in the x-y plane as compared to the z-axis and thus adds more uncertainty in the x-y plane. The positioning performance saturates with increasing numbers of pairs since the addition of low energy pairs does not contribute much to performance improvement.



(a)

(b)

Figure 3.9: Contour plots of CRLB [cm] against spatial variation in the room for an object of fixed height ($z_o = 1.6$ m) for (a) $\beta = 0.25$, (b) $\beta = 1$.

3.6.5 CRLB variation over the room

The CRLB over a spatial grid is evaluated by moving an object of fixed height across the room. All the source-receiver pairs are used to evaluate CRLB. The discretization is done along x and y axes at a resolution of 10 cm and the CRLB is computed at each point. Figure 3.9 presents contour plots of CRLB by using different brightness levels of sources i.e., varying β. It can be observed from the plots that the CRLB is improved by increasing the brightness in room. This is expected because changing brightness of the sources means changing SNR of the received OIR. Notice also that the CRLB is the highest along edges of the room. This is because when the object is around those locations then only a few source-receiver pairs are visible and able to contribute to the positioning estimate. It is also observed that the CRLB has small fluctuations within the space between the luminaires and the receivers. This effect is most prominent at low dimming levels, i.e., $\beta = 0.25$, and is largely due to the fact that limited FOV of the receivers excludes certain pairs from the position estimate.

3.7 Conclusions

Passive indoor positioning using luminaires and receivers in a room is proposed where positioning estimates are derived using reflected energy from an object in the room. Unlike earlier active positioning approaches, our passive positioning approach does not require direct user intervention.

In particular, we model the reflected object impulse response (OIR) between different source-receiver pairs using exponential-integrating sphere model. The model is represented as a function of the object's position as well as source and receiver parameters, and the physical geometry of the room. The OIR measurements can be obtained from each source-receiver pair in a short interval of time by using a coordinating backbone network. Both an ML estimator as well as an analytical CRLB expression are derived.

The results show that performance is improved both by increasing the number of collected source-receiver signals as well as the brightness of the luminaires. Positioning performance is shown to be most sensitive to the dimming level which controls the overall SNR of the system. The performance is also improved by increasing the number of source-receiver pairs measured, however, it saturates quickly. This is consistent with the fact that performance is dominated by several high energy source-receiver pairs. Thus, at high brightness satisfactory performance can be obtained with less complexity by choosing the optimum number of OIR measurements to use in the estimation. When the user dims the lights, similar performance is possible, at greater complexity, by collecting more OIR measurements from a greater number of source-receiver pairs. With typical in room parameters, RMSE of less that 10 cm was found both through simulation and in the developed lower bounds on positioning performance.

Future work includes extending the formulation to consider multiple objects in the room. In addition, the impact of tilt and misalignment in the sources and receivers as well as synchronization errors will also be considered in the future to reflect realistic physical impairments.

3.A Partial Derivatives

The partial derivatives needed to compute the CRLB given in Section 3.4 are rather involved and the procedure to obtain them is described in this appendix. The partial derivative in the first term of (3.4.7) with respect to x_o is given as

$$\frac{\partial}{\partial x_o}\left\{t_{s_i r_j}\left(\mathbf{x}_o\right)\right\} = \left(\frac{x_o - x_{s_i}}{c\left\|\mathbf{x}_o - \mathbf{x}_i^{(s)}\right\|_2} - \frac{x_{r_j} - x_o}{c\left\|\mathbf{x}_j^{(r)} - \mathbf{x}_o\right\|_2}\right). \tag{3.A.1}$$

The derivative of the second term in (3.4.7) can be expanded using the product rule of derivatives as shown in (3.A.2). The partial derivative of $\tau_{s_i r_j}\left(\mathbf{x}_o\right)$ in (3.A.2) can be further expanded as shown in (3.A.3). As an example, (3.A.4) presents the partial derivative of time delay from corner a_1 of $\mathcal{S}_o$ elaborated in Figure 3.3. The derivatives for other corners can be derived in the similar manner.

$$\begin{aligned}\frac{\partial}{\partial x_o}\left\{\eta_{s_i r_j}\left(\mathbf{x}_o\right) e^{-\frac{\left(t-t_{s_i r_j}(\mathbf{x}_o)\right)}{\tau_{s_i r_j}(\mathbf{x}_o)}}\right\} =& \eta_{s_i r_j}\left(\mathbf{x}_o\right) e^{-\frac{\left(t-t_{s_i r_j}(\mathbf{x}_o)\right)}{\tau_{s_i r_j}(\mathbf{x}_o)}} \\ & \left(\frac{\tau_{s_i r_j}\left(\mathbf{x}_o\right)\frac{\partial t_{s_i r_j}(\mathbf{x}_o)}{\partial x_o} - \left(t_{s_i r_j}\left(\mathbf{x}_o\right) - t\right)\frac{\partial \tau_{s_i r_j}(\mathbf{x}_o)}{\partial x_o}}{\left(\tau_{s_i r_j}\left(\mathbf{x}_o\right)\right)^2}\right) \\ & + \eta_{s_i r_j}\left(\mathbf{x}_o\right) e^{-\frac{\left(t-t_{s_i r_j}(\mathbf{x}_o)\right)}{\tau_{s_i r_j}(\mathbf{x}_o)}} \\ & \left(\frac{4\left(x_{r_j} - x_o\right)}{\left\|\mathbf{x}_j^{(r)} - \mathbf{x}_o\right\|_2^2} - \frac{(m+3)\left(x_o - x_{s_i}\right)}{\left\|\mathbf{x}_o - \mathbf{x}_i^{(s)}\right\|_2^2}\right)\end{aligned} \tag{3.A.2}$$

$$\begin{aligned}\frac{\partial \tau_{s_i r_j}\left(\mathbf{x}_o\right)}{\partial x_o} =& \frac{1}{4\tau_{s_i r_j}\left(\mathbf{x}_o\right)} \\ & \times\left\{\left(t_{s_i r_j}^{(a_1)}\left(\mathbf{x}_o\right) - t_{s_i r_j}\left(\mathbf{x}_o\right)\right)\left(\frac{\partial t_{s_i r_j}^{(a_1)}\left(\mathbf{x}_o\right)}{\partial x_o} - \frac{\partial t_{s_i r_j}\left(\mathbf{x}_o\right)}{\partial x_o}\right)\right. \\ & + \left(t_{s_i r_j}^{(a_2)}\left(\mathbf{x}_o\right) - t_{s_i r_j}\left(\mathbf{x}_o\right)\right)\left(\frac{\partial t_{s_i r_j}^{(a_2)}\left(\mathbf{x}_o\right)}{\partial x_o} - \frac{\partial t_{s_i r_j}\left(\mathbf{x}_o\right)}{\partial x_o}\right) \\ & + \left(t_{s_i r_j}^{(a_3)}\left(\mathbf{x}_o\right) - t_{s_i r_j}\left(\mathbf{x}_o\right)\right)\left(\frac{\partial t_{s_i r_j}^{(a_3)}\left(\mathbf{x}_o\right)}{\partial x_o} - \frac{\partial t_{s_i r_j}\left(\mathbf{x}_o\right)}{\partial x_o}\right) \\ & \left. + \left(t_{s_i r_j}^{(a_4)}\left(\mathbf{x}_o\right) - t_{s_i r_j}\left(\mathbf{x}_o\right)\right)\left(\frac{\partial t_{s_i r_j}^{(a_4)}\left(\mathbf{x}_o\right)}{\partial x_o} - \frac{\partial t_{s_i r_j}\left(\mathbf{x}_o\right)}{\partial x_o}\right)\right\}\end{aligned} \tag{3.A.3}$$

$$\begin{aligned}\frac{\partial t_{s_i r_j}^{(a_1)}\left(\mathbf{x}_o\right)}{\partial x_o} =& \frac{c^{-1}\left(\left(x_o - \Delta x\right) - x_{s_i}\right)}{\sqrt{\left(\left(x_o - \Delta x\right) - x_{s_i}\right)^2 + \left(\left(y_o + \Delta y\right) - y_{s_i}\right)^2 + \left(z_o - z_{s_i}\right)^2}} \\ & - \frac{c^{-1}\left(x_{r_j} - \left(x_o - \Delta x\right)\right)}{\sqrt{\left(x_{r_j} - \left(x_o - \Delta x\right)\right)^2 + \left(y_{r_j} - \left(y_o + \Delta y\right)\right)^2 + \left(z_{r_j} - z_o\right)^2}}\end{aligned} \tag{3.A.4}$$

The partial derivatives with respect to y_o can be derived in the same fashion since there exists symmetry in x and y coordinates.

Consider the partial derivative with respect to z_o in (3.4.7) (i.e., $k = 3$). The partial derivative in first term of (3.4.7) with respect to z_o is similar to (3.A.1). The partial derivative of second term of (3.4.7) with respect to z_o can be expanded similarly using the product rule of derivatives and is given in (3.A.5). The partial derivative $\tau_{s_i r_j}(\mathbf{x}_o)$ with respect to z_o can be expanded in a similar fashion to (3.A.3). As was done for x and y coordinates, the partial derivative of time delay from corner a_1 of the object top with respect to z_o is given in (3.A.6). The partial derivatives for other corners can be derived similarly.

$$\begin{aligned}\frac{\partial}{\partial z_o}\left\{\eta_{s_i r_j}(\mathbf{x}_o)\, e^{-\frac{\left(t-t_{s_i r_j}(\mathbf{x}_o)\right)}{\tau_{s_i r_j}(\mathbf{x}_o)}}\right\} &= \eta_{s_i r_j}(\mathbf{x}_o)\, e^{-\frac{\left(t-t_{s_i r_j}(\mathbf{x}_o)\right)}{\tau_{s_i r_j}(\mathbf{x}_o)}} \\ &\left(\frac{\tau_{s_i r_j}(\mathbf{x}_o)\frac{\partial t_{s_i r_j}(\mathbf{x}_o)}{\partial z_o} - \left(t_{s_i r_j}(\mathbf{x}_o) - t\right)\frac{\partial \tau_{s_i r_j}(\mathbf{x}_o)}{\partial z_o}}{\left(\tau_{s_i r_j}(\mathbf{x}_o)\right)^2}\right) \\ &+ \eta_{s_i r_j}(\mathbf{x}_o)\, e^{-\frac{\left(t-t_{s_i r_j}(\mathbf{x}_o)\right)}{\tau_{s_i r_j}(\mathbf{x}_o)}} \\ &\left(-\frac{2}{z_{r_j} - z_o} - \frac{(m+1)}{z_{s_i} - z_o} + \frac{4\left(z_{r_j} - z_o\right)}{\left\|\mathbf{x}_j^{(r)} - \mathbf{x}_o\right\|_2^2} - \frac{(m+3)\left(z_o - z_{s_i}\right)}{\left\|\mathbf{x}_o - \mathbf{x}_i^{(s)}\right\|_2^2}\right)\end{aligned} \tag{3.A.5}$$

$$\frac{\partial t_{s_i r_j}^{(a_1)}\left(\mathbf{x}_o\right)}{\partial z_o} = \frac{c^{-1}\left(z_o - z_{s_i}\right)}{\sqrt{\left(\left(x_o - \Delta x\right) - x_{s_i}\right)^2 + \left(\left(y_o + \Delta y\right) - y_{s_i}\right)^2 + \left(z_o - z_{s_i}\right)^2}} - \frac{c^{-1}\left(z_{r_j} - z_o\right)}{\sqrt{\left(x_{r_j} - \left(x_o - \Delta x\right)\right)^2 + \left(y_{r_j} - \left(y_o + \Delta y\right)\right)^2 + \left(z_{r_j} - z_o\right)^2}} \qquad (3.A.6)$$

Acknowledgment

The authors would like to thank Drs. R. Tharmarasa and T.N. Davidson for their helpful discussions in the formulation of this work.

Chapter 4

Passive Indoor Visible Light Positioning System using Deep Learning

IEEE Copyright Notice

The research work presented in this chapter is accepted for publication in the following journal article [89].

- **K. Majeed** and S. Hranilovic, "Passive Indoor Visible Light Positioning System using Deep Learning," in *IEEE Internet of Things Journal*, doi: 10.1109/JIOT.2021.3072201.

The formatting of the manuscript presented in this chapter is changed in order to follow the thesis format e.g., equation numbers, citations, figures, etc.

In contrast to single-bounce model in Chapters 2 and 3, a realistic room model is used in this chapter in order to characterize the performance of the passive indoor visible light positioning systems. The arrangement of sources and receivers and impulse response (IR) acquisition process is the same as described in Chapters 1, 2, and 3.

The sources are modeled to have wavelength-dependent flux distribution and receivers are assumed to have certain field-of-view and fixed area. The room also contains furniture as compared to the single-bounce model used in Chapters 2 and 3. Furthermore, the coatings used to model the optical reflectivity of surfaces in the room e.g., floor, ceiling, walls, and furniture have reflection characteristics that depend on wavelength of light source. The localization object (LO) is modeled as a cuboid of fixed height with a head-like shape at the top in order to represent the human head.

The IR measurements corresponding to a particular location of the LO are obtained between all source-receiver pairs affixed in the room and are collected in a set. In order to use deep learning, the sets of IRs between all source-receiver pairs are collected when the LO is placed at uniformly distributed positions in the room. The collected sets of IRs are represented as features and divided into training and validation data sets with corresponding positions. The training and validation data are input to the neural network architecture after preprocessing in order to train the network. Once the network is trained it can be used to estimate position of the LO from a set of IR measurements when the LO is present at a random location.

The proposed method shows root-mean-square error around 30 cm by using luminaires at full brightness and around 80 cm when the network is trained by using only 10 randomly selected locations in the room.

In order to represent features of IR measurements, a binning process is used instead of filtering or peak detection. It is important to note that the peak corresponding to the LO is quite small in amplitude and time spread as compared to rest of the IR waveform. Peak detection becomes complex due to noise sensitivity at the receiver. On the other hand, the binning process limits filtering of the noise to the respective bins and aids in highlighting features corresponding to the LO especially at small bin sizes.

Secondly, the IR measurements obtained at the receiver contain noise that is a combination of different sources of noise e.g., thermal noise, shot noise, etc. The noise at the receiver is modeled as additive white Gaussian noise and is considered independent between the receivers. The IR measurements contain noise per sample according to the assumed noise variance and the time binning changes the value of SNR in the respective bin.

The realistic room scenario is modeled in Zemax® OpticStudio [132] which is a commercial ray tracing software. The software uses global illumination algorithms [133] and recursive or Monte Carlo ray tracing [106, 134] approaches in order to compute intensity of the traced ray segments. Recent work in [135] verifies the closeness of experimental channel measurements with the simulated channel model obtained from Zemax®.

4.1 Abstract

A passive indoor visible light positioning (VLP) system is proposed that does not require active participation from the user and is suitable for IoT sensor networks. This approach does not require a line-of-sight path and measures the impulse response (IR) between sources and receivers installed in the room. The presence of an object of interest (OI), i.e. a person to be localized, disrupts the IRs among the source-receiver pairs which can be related to its position. A deep learning framework is developed which learns the relationship between changes in sets of IRs and the OI position through a set of training data obtained by placing the OI at random locations in the room. This approach shows that the OI can be localized using a very limited set of training data under a wide range of illumination levels. In order to represent a realistic scenario, a room with furniture is modelled in optical system design software. The ray trace information of the modelled room is used to construct IR measurements among different source-receiver pairs that include multi-order reflections. The results show that localization performance is crucially related to signal-to-noise ratio and number of training data points used in the learning process. A root-mean square error (RMSE) near 30 cm is possible in the case of high SNR and a large training set. However, even with a very limited training set and over a range of dimming levels, RMSEs of near 80 cm were obtained without the need for explicit user involvement.

4.2 Introduction

The ubiquity of light-emitting diodes (LEDs) in indoor areas has tremendously enabled the development of visible light communication (VLC) systems [45] and visible light positioning systems (VLP) [46]. These LED luminaires are capable of providing communication and illumination simultaneously in an indoor area to support a wide variety of IoT applications. Such VLP systems provide better accuracy, security and privacy as compared to their radio frequency (RF)-based counterparts and do not require extra hardware. The development of VLP has led in turn to the development of various algorithms and techniques to leverage indoor luminaires for localization [37, 113, 114, 136, 101, 59, 119, 137, 117, 115, 60, 138, 74, 139, 140, 141, 111, 51, 142, 143, 144, 129, 70, 46, 145, 146, 108, 65, 52, 47]. These approaches can be broadly classified into the following categories: proximity-based, scene analysis-based and trilateration/triangulation-based techniques.

Proximity-based techniques [37, 113, 114] provide a coarse position estimate of a receiver. In order to approximate the location of a receiver, a unique code is transmitted by a source that is identified at the detector or camera in the receiver. The nearest identified source in the vicinity of the receiver thus provides its position estimate. *Scene analysis*-based techniques [136, 101, 59, 119, 137, 117, 115, 60] use received signal strength (RSS) from multiple sources at the receiver and typically provide better accuracy than proximity-based techniques. The RSSs can be related to the distance of a receiver from a set of sources in order to provide a position estimate of the receiver. *Triangulation/trilateration*-based techniques [138, 139, 140, 141, 111], on the other hand, measure time of arrival (TOA), time difference of arrival (TDOA), or angle of arrival (AOA) of the received signals in order to estimate receiver's position. These techniques require sensitive hardware at the receiver and also need perfect synchronization between the sources and receivers.

The aforementioned techniques fall under the category of *active* localization where a user is equipped with a receiver. These techniques require input from the user e.g., holding a receiver and initiating a localization request, hence the name, *active*. There is also a direct line-of-sight (LOS) path between the source and receiver and thus the localization algorithms rely highly on the power received from the LOS component. In contrast, *passive* VLP approaches do not require explicit intervention of the user and often rely on both LOS as well as diffuse components [87, 88]. The localization process is typically based on existing system infrastructure where measurements are obtained between the luminaires and receivers installed in the room rather than using user equipped receivers/sources.

A review of some passive VLP techniques can be found in [53]. Most of these works sense location of an *object of interest* (OI) by studying effects of received power due to the shadows formed by the OI in the room. Some of these works use a special arrangement of luminaires and receivers e.g., [147]. A related passive approach can be found in [76], where position estimates are provided by using power measurements in an outdoor scenario. Another passive approach is proposed in [148, 79], where the blockage of LOS signals between sources and receivers are used to detect the location of the OI. The proposed system in [148] uses transceivers affixed on the ceiling and multiple receivers (e.g., phones, autonomous devices, wearable devices, etc.) that are moving or stationary. The method in [80] uses a similar system topology that includes multi-order reflections along with the derivation of Cramér-Rao lower bound (CRLB) on fingerprinting-based positioning. The aforementioned techniques either require hardware modifications or special arrangement of luminaires and receivers in an indoor area.

In this paper, a proof-of-concept of passive VLP system is designed and tested in a realistic indoor scenario where position estimates are obtained using a deep learning framework. Though similar in architecture to our earlier theoretical work [87, 88], this paper presents a first system design for indoor passive VLP systems. In [88], a CRLB on passive VLP in an idealistic scenario is derived, where the room is considered empty and the acquired impulse responses (IRs) contain first-order reflections only. The approach in [87] requires fingerprinting map of IRs in the indoor area, which uses exhaustive search to estimate position and also increases storage memory requirements. This paper extends our earlier work by considering a complex realistic scenario in which a room containing furniture is modelled and the IRs among the source-receiver pairs include multi-order reflections. The different entities and surfaces in the room are represented with characteristics based on measured spectral data. This commercial simulator allows considerable flexibility to accurately model a wide variety of indoor scenarios and the focus of this paper is on the development of algorithms to passively localize an indoor user with diffuse reflections.

This work demonstrates that the passive localization approach, when coupled with a deep learning framework, is able to localize the OI under a wide range of illumination levels and with limited training. Though this proof-of-concept work considers only a single OI in the room for localization, the deep-learning framework has potential to be extended. One of the crucial applications of passive VLP is to construct an IoT sensor network to monitor patients in hospital or at home while simultaneously preserving their privacy.

The following notations are used in this paper: scalar by x, vector by $\mathbf{x}$, matrix by $\mathbf{X}$, set by $\mathcal{X}$, j^{th} element of a vector (with subscript i) $\mathbf{x}_i$ by $x_j^{(i)}$, cardinality of a set $\mathcal{X}$ by $|\mathcal{X}|$. The transpose operator is denoted by $[\cdot]^T$ and the Euclidean norm by

$\| \cdot \|_2$. All vectors in this paper are column vectors.

Sec. 4.3 describes the passive indoor localization system, a realistic modelling of the indoor environment, and the measurement of IRs. A deep learning framework for position estimation is explained in Sec. 4.4 and the performance of the proposed system is evaluated in Sec. 4.5. Finally, the paper is concluded in Sec. 4.6.

4.3 System Model

This section describes the passive localization scenario with a proposed system architecture and realistic modeling of a room in Zemax® OpticStudio [132]. The process of constructing IRs by using the data gathered from Zemax® is also explained along with the measurement model for IR acquisition.

4.3.1 Passive Localization Scenario

In a passive localization scenario there is not necessarily a direct LOS path between a source-receiver pair and thus the light rays are detected at the receiver after potentially undergoing multi-order reflections. The order here refers to the number of times a ray hits any surface or entity in the room (e.g., walls, floor, ceiling, furniture, OI, etc.) before they are detected at the receiver. An example of a passive scenario is depicted in Fig. 4.1, which shows cross section of the room containing luminaires and receivers co-located on the ceiling. Notice that this definition of passive localization does not impose a strict criterion on the location of sources and receivers. The co-location of sources and receivers is considered here given the convenience for implementation since most indoor areas are equipped with symmetrical arrangement

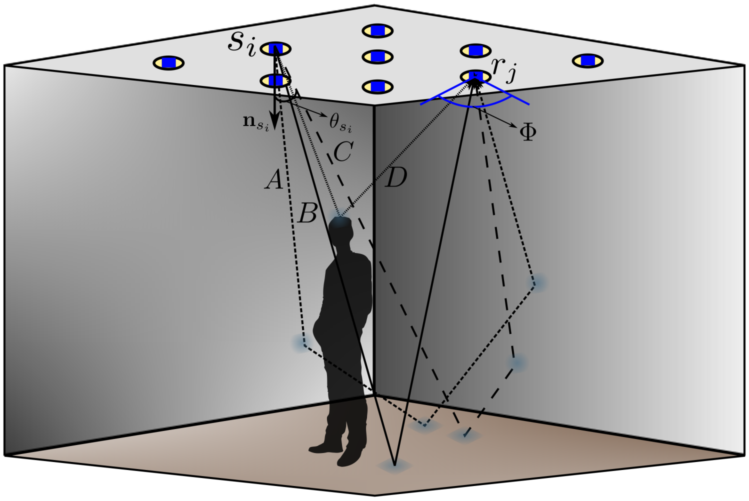


Figure 4.1: A room containing transceivers (luminaires and receivers) co-located on the ceiling. The light rays A, B, C and D after originating from source s_i and undergoing multi-order reflections are detected at receiver r_j. Notice that the path of D is influenced by presence of a person (clipart silhouette reproduced from [55]).

of luminaires on the ceiling. The proposed passive localization approach does channel sounding between each of the sources and receivers that is similar to underwater acoustic channel sounding [103]. In the particular case of co-located sources and receivers, shown in Fig. 4.1, notice that only non-LOS components of the IR with multi-order reflections are detected at the receiver. For instance, consider the light rays A, B and C originating from source s_i that undergo 3^{rd}, 1^{st} and 2^{nd}-order reflections respectively, which are indicated by color spot at each point of incidence on the surface before being detected at the receiver r_j. Additionally, the path of ray D from source to receiver is influenced by the position of the OI (i.e., a person in the room). The collection of detected rays are then used to construct the IR between source s_i and receiver r_j.

The luminaires are considered to have a Lambertian radiation pattern [149, 45]

$R(\theta)$ in Watts/steradian given as

$$R(\theta) = \frac{m+1}{2\pi}(\alpha P_s)\cos^m(\theta) \tag{4.3.1}$$

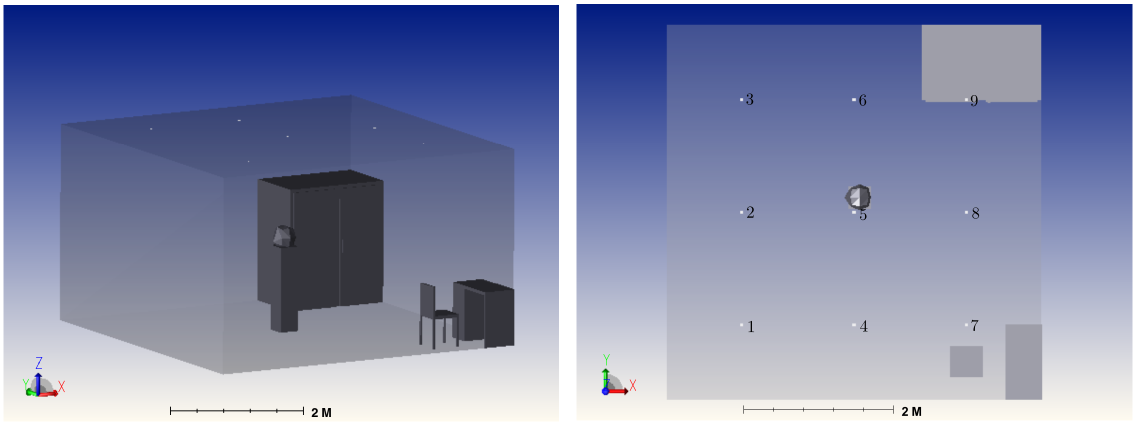
where θ denotes the angle formed between a ray emerging from the source and normal to the source, $\alpha \in (0, 1]$ denotes a dimming parameter that controls brightness of the luminaire, P_s is the maximum power emitted by the source in Watts, and m denotes Lambertian index of the source. Notice that the radiation pattern in (4.3.1) has uniaxial symmetry.

The receivers are assumed to have a fixed area of A_r, field-of-view (FOV) Φ, and responsivity γ.

Define $h_{s_i r_j}(t)$ as the IR between source s_i and receiver r_j. The IRs among all combinations of N_s sources $\{s_i | i = 1, 2, \cdots, N_s\}$ and N_r receivers $\{r_j | j = 1, 2, \cdots, N_r\}$ are accumulated in set $\mathcal{H}$ with cardinality $|\mathcal{H}| = N_s \times N_r$ as

$$\mathcal{H} = \left\{ \begin{array}{c} h_{s_1 r_1}(t), h_{s_1 r_2}(t), \cdots, h_{s_1 r_{N_r}}(t) \\ h_{s_2 r_1}(t), h_{s_2 r_2}(t), \cdots, h_{s_2 r_{N_r}}(t) \\ \vdots \\ h_{s_{N_s} r_1}(t), h_{s_{N_s} r_2}(t), \cdots, h_{s_{N_s} r_{N_r}}(t) \end{array} \right\} \tag{4.3.2}$$

In practice, a sampled version of the $h_{s_i r_j}(t)$ that are limited to a certain time extent are used for OI position estimation as described in Sec. 4.3.3.



(a) (b)

Figure 4.2: A realistic model of the room in Zemax® that contains source-receiver pairs, furniture and OI. The source-receiver pairs are affixed on the ceiling as indicated by white dots and are co-located. The furniture is located towards edge of the room along positive x-axis. An example placement of the OI at (2.57 m, 2.70 m, 1.6 m) is also shown. The dimensions of the room are 5m × 5 m × 3 m. A graphic illustration of the room: (a) 3D view and (b) top view.

4.3.2 Room Modeling using Realistic Parameters

The optical system design software Zemax® [132] is used to model a room. The non-sequential ray tracing mode of the software is used to obtain ray trace information, where the rays originating from a source hit surfaces or objects depending on their locations in the room. The generated rays can hit small portions of an object or a surface multiple times while being traced. The modeled room is shown in Fig. 4.2, where the furniture is located towards edge of the room along positive x-axis. The figure shows both 3D and top view of the modeled room. An example placement of OI near center of the room is also shown for illustrative purposes. The room is modeled in two steps as described in the following.

In the first step, the furniture is modeled as CAD objects in Blender [150] that includes a closet, table and chair as shown in Fig. 4.2. Typical dimensions of the

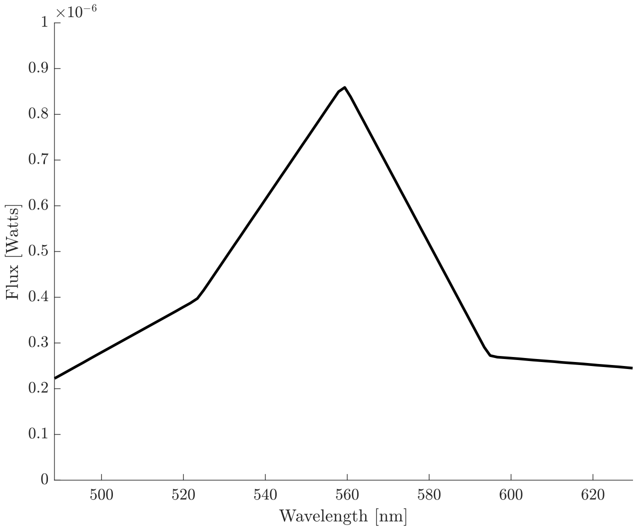


Figure 4.3: Flux variation of the source against the visible light wavelengths used for ray tracing in Zemax®. The graph is obtained by using flux versus wavelength analysis in Zemax® [132].

furniture are listed in Table 4.1. The OI is modeled as a cuboid with head-like shape at its top to resemble a human head. In the second step, the room environment is set up in Zemax® and the CAD objects are imported and placed inside the room. The source-receiver pairs are affixed on the ceiling as described in Section 4.3.1 and are co-located.

The sources are modeled by using the *source rectangle* object, which provides flexibility to control the Lambertian index m of the source. The source is considered photopic that has variable flux distribution against the visible light wavelengths as depicted in Fig. 4.3 [132]. The receivers are modelled using the *detector rectangle* object in Zemax®, which provides flexibility to control the size and FOV Φ of the detector. Nominal room temperature and pressure are set as given in Table 4.1. All surfaces in the room are defined by reflectivity of typical materials. Pine wood is defined to coat the furniture and plaster is defined to coat the walls. The reflectivity of

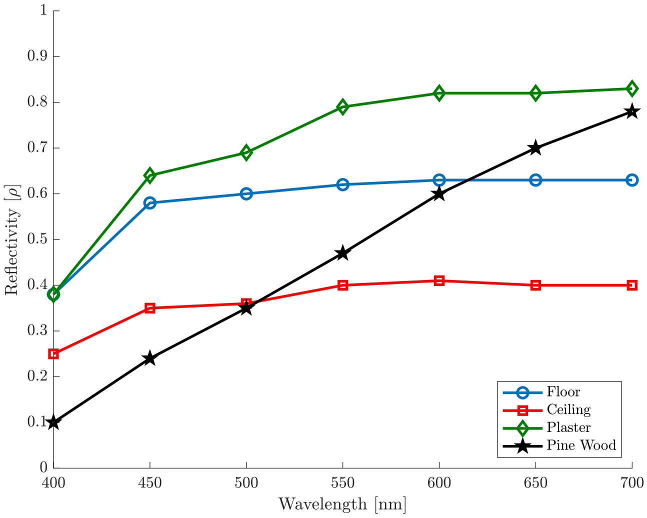


Figure 4.4: Reflectivity values of real materials [93, 95, 94] against the visible light wavelengths that are used to coat room surfaces and furniture used in Zemax®.

the coating materials has measured wavelength-dependent characteristics [93, 95, 94] as shown in Fig. 4.4. Moreover, the surfaces of different entities in the room are modelled as diffuse reflectors.

4.3.3 Construction of Impulse Response Waveform

The ray trace of the modeled indoor environment in Zemax® produces a database file that contains important information (e.g., path length, intensity, hit object number, level of segments, etc.) about the rays detected at the receiver. The values of these parameters are essential to construct the IRs between all source-receiver pairs. The path length and intensity of the rays detected at the receiver corresponding to a certain source-receiver pair are used to construct the IR between that pair. For example,

$h_{s_i r_j}(t)$ can be approximated from the ray trace data obtained from simulation as

$$h_{s_i r_j}(t) \approx \sum_{k=1}^{N_o} p_k \delta(t - \tau_k) \tag{4.3.3}$$

where p_k denotes the intensity of last segment of the k^{th} ray detected at the receiver r_j, $\tau_k = \frac{d_k}{c}$ denotes the time taken by the k^{th} ray to travel a total path length d_k from source s_i to receiver r_j after multi-order reflections, N_o denotes the total number of rays detected at the receiver r_j in the simulation, and c is the speed of light. This process is repeated to construct the IRs between all source and receiver pairs. The constructed IRs are set to span the range $t = \{0, t_{\max}\}$, where $t_{\max}$ denotes the maximum time chosen for IR acquisition. The choice of $t_{\max}$ depends on both the geometry of the room as well as maximum storage capacity of database.

As an example, consider $h_{s_5 r_5}(t)$ the IR between source s_5 and receiver r_5 as shown in Fig. 4.5 when the OI is located near center of the modeled room (see Fig. 4.2). The figure shows the constructed IR without and with the presence of OI in the room. It can be observed from the figure that a small peak occurs around 10 ns when the OI can be located in vicinity of receiver r_5. The comparison of IRs shows that the presence of the OI alters the IRs between source-receiver pairs, where the observed changes are considerable in the first-order reflection and relatively smaller in higher-order reflections. Similarly, the IRs among other source-receiver pairs also show small changes in them according to the location of OI. In Sec. 4.4, these changes in the IRs among the source-receiver pairs are leveraged in a deep learning framework to estimate the position of OI in the room.

It is important to note that the proposed system considers an OI of fixed height. This application scenario can arise, for example, in cases where monitoring of patients

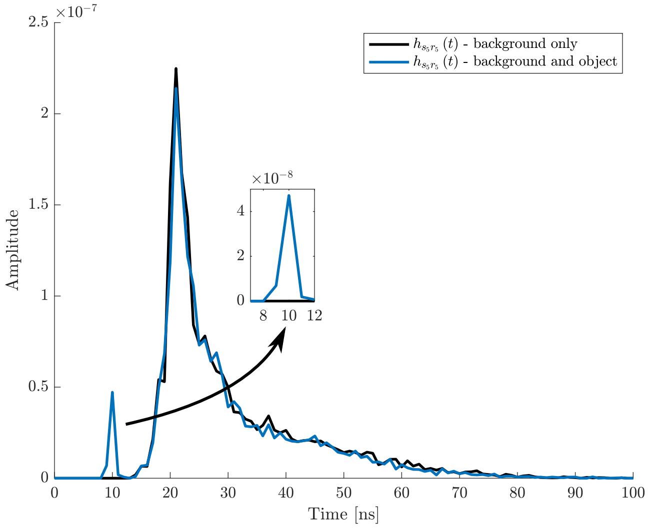


Figure 4.5: Impulse response $h_{s_5 r_5}(t)$ between source s_5 and receiver r_5 when the OI is located at (2.57 m, 2.70 m, 1.6 m) as illustrated in Fig. 4.2. The values of different parameters used are listed in Table 4.1.

in a hospital room is required. The observation of IR in Fig. 4.5 shows that the initial peak occurring in IR waveform around 10 ns is due primarily from reflections from the top surface of the OI while the peak at 25 ns is due to reflections from the floor. Reflections from side surfaces of the OI appear later in time and have considerably less energy as compared to the top surface reflection pulse. Furthermore, multi-order reflections from the OI are very small in energy as compared to the first bounce from the top surface and are very small compared to the dominant reflection from the floor. Therefore, the position of OI is crucially related to the first peak in the IR corresponding to the first bounce from the top of OI, which plays an important part in the localization using the deep learning framework.

4.3.4 Measurement Model of Impulse Response Waveform

In order to acquire the IRs, it is assumed that the light sources and receivers are arranged as a network at the back end that is controlled from a central entity. The measurement process of IRs is achieved by setting up a common time reference among all the receivers and a single source is turned on at a time. The individual light sources are turned on and off at an imperceptibly fast rate so that light flicker does not impact on the illumination performance of the luminaires given that the main role of light sources in the indoor environment is to provide illumination. The IRs among all source-receiver pairs can be accumulated in a fraction of a second since the length of a typical IR in the room is on the order of 10's of ns.

Define $g_{s_i r_j}(t)$ the IR measured between source s_i and receiver r_j modelled as

$$g_{s_i r_j}(t) = h_{s_i r_j}(t) + n_{s_i r_j}(t) \tag{4.3.4}$$

where $n_{s_i r_j}(t)$ is additive white Gaussian noise (AWGN) between source s_i and receiver r_j with mean zero and variance σ^2. The noise $n_{s_i r_j}(t)$ is considered independent among all source-receiver pairs. The measured IRs among all source-receiver pairs can be accumulated in set $\mathcal{G}$ as

$$\mathcal{G} = \left\{ \begin{array}{c} g_{s_1 r_1}(t), g_{s_1 r_2}(t), \cdots, g_{s_1 r_{N_r}}(t) \\ g_{s_2 r_1}(t), g_{s_2 r_2}(t), \cdots, g_{s_2 r_{N_r}}(t) \\ \vdots \\ g_{s_{N_s} r_1}(t), g_{s_{N_s} r_2}(t), \cdots, g_{s_{N_s} r_{N_r}}(t) \end{array} \right\} \tag{4.3.5}$$

where cardinality of $\mathcal{G}$ is same as $\mathcal{H}$ i.e., $|\mathcal{G}| = N_s \times N_r$. In other words, the set $\mathcal{G}$ is a noisy version of $\mathcal{H}$. The noise in the measurements in $\mathcal{G}$ can be mitigated by averaging repeated IR measurements [151] between each source-receiver pair as discussed in Sec. 4.5 at the expense of greater latency in producing a location estimate.

The task in passive VLP is to localize the user using a measurement of $\mathcal{G}$ while the user is at a fixed location. As described in Sec. 4.4, multiple sets ($\mathcal{G}$'s) of IRs among all source-receiver pairs are obtained by placing the OI at randomly selected positions in the room to produce a training set to learn the relationship between IRs and positions.

4.3.5 Modelling of Indoor VLC Environments

The indoor environment in this work is modelled using the commercial ray tracing software tool Zemax® [132]. This software tool allows for the control of source, receiver, and coating material parameters along with multi-order reflections inside of the room. In the context of indoor optical wireless systems, ray tracing approaches are the most popular modelling tools used in the literature. For example, the seminal approach of Barry *et al.* in [106] is a recursive approach to simulate IRs and model multi-order reflections in indoor environment. The authors of [106] demonstrate a close correspondence between the simulated and measured IRs in indoor environments.

Apart from the recursive approach of [106], several methods [152, 153, 154, 155, 135] have employed Zemax® Opticstudio in order to model a realistic indoor environments for VLC systems. Recently, [135] specifically compares the indoor optical IR between those simulated in Zemax® and experimental measurements. The reported mean-square error between the experimental and simulated channel model is within 2% in the presence of multi-order reflections. Given this earlier work verifying the accuracy of Zemax® in characterizing IRs in indoor VLC environments, in Sec. 4.5 numerical results on our passive localization approach are provided based on ray tracing models as well as measured optical parameters for room surfaces.

4.4 Deep Learning Framework for Position Estimation

In this section, a deep learning framework is employed that learns the relationships between the changes in the IRs between source-receiver pairs due to the OI in order to infer position. The network does not directly learn the position of OI but rather the relationship between changes in IRs due to the presence of the OI which is used to infer its position. The IRs are first preprocessed before feeding them to the neural network. The proposed framework uses a feed-forward neural network (FNN) architecture with multiple hidden layers [96]. Fig. 4.6 shows the layout of complete deep learning framework employed for position estimation.

4.4.1 Feature Representation of Impulse Response Waveform

The IRs among all source-receiver pairs are divided into $N_b = \left\lceil \frac{t_{\max}}{t_b} \right\rceil$ time bins each i.e.,

$$\mathbb{T}_n = \begin{cases} [(n-1)\, t_b, n t_b), & n = 1, 2, \cdots, N_b - 1 \\ [(N_b - 1)\, t_b, t_{\max}), & n = N_b \end{cases} \tag{4.4.1}$$

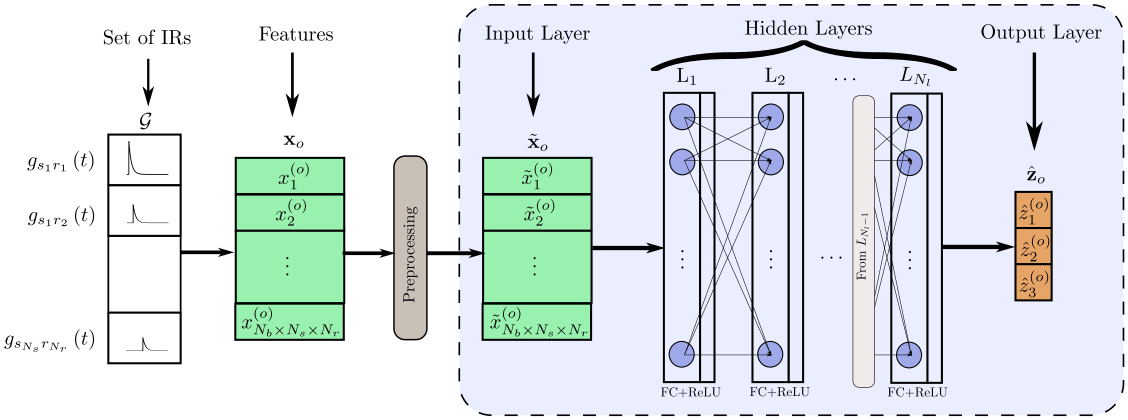


Figure 4.6: Deep learning framework showing set of measured IRs $\mathcal{G}$, feature-based vector $\mathbf{x}$ and the FNN architecture [96]. The FNN contains $N_b \times N_s \times N_r$ dimensional input layer of preprocessed features, which is followed by N_l hidden layers with each layer composed of fully connected (FC) and ReLU layers. The last layer is N_z dimensional output layer estimating the position of the OI.

where t_b is the duration chosen for the time bin and $\mathbb{T}_n$ denotes the n^{th} time bin. It is important to note that when the ratio $\frac{t_{\max}}{t_b}$ is an integer then all the N_b time bins have equal size, t_b, and when the ratio is a non-integer then the first $N_b - 1$ time bins have equal size, t_b, and the last N_b^{th} time bin has smaller size, i.e., $t_{\max} - (N_b - 1)t_b$.

As an example, consider the measured IR $g_{s_5r_5}(t)$ between source s_5 and receiver r_5 with N_b time bins as shown in Fig. 4.7 (notice that this corresponds to a noisy version of the IR in Fig. 4.5). Recall that $g_{s_5r_5}(t)$ corresponds to the OI located near center of the room as illustrated in Fig. 4.2. The portion of the IR in each time bin is integrated to represent a feature which results in N_b features in total for single IR measurement.

The selection of t_b affects the performance of the system especially at low dimming level of luminaires as discussed in Sec. 4.5. Moreover, since the receiver typically has a band-limited response thus representing IRs with large time bins t_b (or alternately

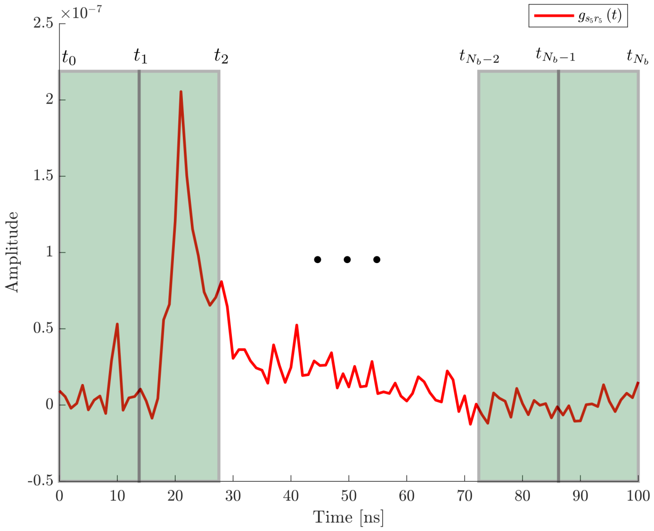


Figure 4.7: An example of measured IR $g_{s_5 r_5}(t)$ between source s_5 and receiver r_5 showing N_b time bins. The OI is located at (2.57 m, 2.70 m, 1.6 m) as shown in Fig. 4.2.

smaller N_b) reduces complexity of IR acquisition in the analog-to-digital (A2D) converter at the receiver.

Consider the measured IR $g^{(o)}_{s_i r_j}(t) \in \mathcal{G}_o$ between source s_i and receiver r_j when the OI is located at o^{th} position $\mathbf{z}_o = [\ z^{(o)}_1 \quad z^{(o)}_2 \quad z^{(o)}_3\ ]^T$, where $z^{(o)}_1$, $z^{(o)}_2$ and $z^{(o)}_3$ are the Cartesian coordinates of the OI. The integral feature $f^{(o)}_{s_i r_j, n}$ of the n^{th} time bin $\mathbb{T}_n$ for $g^{(o)}_{s_i r_j}(t)$ is

$$f^{(o)}_{s_i r_j, n} = \int\limits_{\mathbb{T}_n} g^{(o)}_{s_i r_j}(t)\, dt \tag{4.4.2}$$

Similarly, all of the remaining IRs in $\mathcal{G}_o$ are represented with N_b features each. The features from all IRs in $\mathcal{G}_o$ are concatenated to form the feature-based vector $\mathbf{f}_o$ as

$$\mathbf{f}_o = \left[f^{(o)}_{s_1 r_1, 1} \cdots f^{(o)}_{s_1 r_1, N_b} \cdots f^{(o)}_{s_{N_s} r_{N_r}, 1} \cdots f^{(o)}_{s_{N_s} r_{N_r}, N_b} \right]^T \tag{4.4.3}$$

where the length of $\mathbf{f}_o$ is $N_b \times N_s \times N_r$. It is important to note that (4.4.3) represents features from all source-receiver pairs when the OI is located at position $\mathbf{z}_o$.

In order to collect data for deep learning, the OI is placed at randomly selected positions in the room (shown in Sec. 4.5) and feature-based vectors are obtained at each location by using (4.4.3). The gathered data points are divided into training and validation data sets $\mathcal{X} = \{\mathbf{x}_p | p = 1, 2, \cdots, N_t\}$ and $\mathcal{Y} = \{\mathbf{y}_q | q = 1, 2, \cdots, N_v\}$ respectively. The positions corresponding to the data elements in $\mathcal{X}$ and $\mathcal{Y}$ are collected in data sets $\mathcal{Z}_{\mathcal{X}} = \{\mathbf{z}_p | p = 1, 2, \cdots, N_t\}$ and $\mathcal{Z}_{\mathcal{Y}} = \{z_q | q = 1, 2, \cdots, N_v\}$ respectively. The vector $\mathbf{x}_p \in \mathcal{X}$ is the feature-based vector when the OI is located at $\mathbf{z}_p \in \mathcal{Z}_{\mathcal{X}}$ and can be represented by changing notations in (4.4.3) as

$$\mathbf{x}_p = \left[x^{(p)}_{s_1 r_1, 1} \cdots x^{(p)}_{s_1 r_1, N_b} \cdots x^{(p)}_{s_{N_s} r_{N_r}, 1} \cdots x^{(p)}_{s_{N_s} r_{N_r}, N_b} \right]^T \tag{4.4.4}$$

Similarly, $\mathbf{y}_q \in \mathcal{Y}$ is the feature-based vector when the OI is located at $\mathbf{z}_q \in \mathcal{Z}_{\mathcal{Y}}$ and can be represented as

$$\mathbf{y}_q = \left[y^{(q)}_{s_1 r_1, 1} \cdots y^{(q)}_{s_1 r_1, N_b} \cdots y^{(q)}_{s_{N_s} r_{N_r}, 1} \cdots y^{(q)}_{s_{N_s} r_{N_r}, N_b} \right]^T \tag{4.4.5}$$

It is important to note that the positions corresponding to validation data $\mathcal{Y}$ i.e., $\mathbf{z}_q \in \mathcal{Z}_{\mathcal{Y}}$ are unknown to the deep learning framework since the validation data are only used for performance evaluation. The subscript of elements in vectors $\mathbf{x}_p$ and $\mathbf{y}_q$ are dropped in the following for simplicity as shown in Fig. 4.8 and are chosen equal to the dimensions of feature-based vectors i.e., $N_b \times N_s \times N_r$.

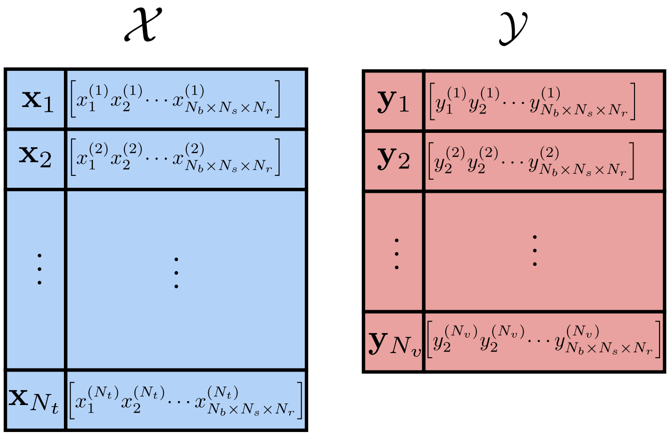


Figure 4.8: Data sets containing feature-based vectors for elements in training data set $\mathcal{X}$ and validation data set $\mathcal{Y}$.

4.4.2 Data Preprocessing

The objective of data preprocessing is to represent all features on a common scale since some features show large variation in values as compared to others. The preprocessing step is a common practice in deep learning since it greatly affects the learning process [96]. The choice of preprocessing method depends on the domain of the problem under consideration. The scaling method is chosen here since the changes observed in IRs among the source-receiver pairs due to OI are related with each other. In order to do the scaling, the maximum value among all features in the training data set $\mathcal{X}$ is computed, which is given as

$$M_{\mathcal{X}} = \max_{\substack{1 \leq i \leq N_b \times N_s \times N_r \\ 1 \leq p \leq N_t}} x_i^{(p)} \tag{4.4.6}$$

The value $M_{\mathcal{X}}$ is then used to normalize data points in both $\mathcal{X}$ and $\mathcal{Y}$. Define $\tilde{\mathbf{x}}_p$ as the p^{th} normalized data point for $\mathbf{x}_p \in \mathcal{X}$ and $\tilde{\mathbf{y}}_q$ the q^{th} normalized data point for

$\mathbf{y}_q \in \mathcal{Y}$, i.e.,

$$\begin{aligned} \tilde{\mathbf{x}}_p &:= \frac{\mathbf{x}_p}{M_{\mathcal{X}}} \\ \tilde{\mathbf{y}}_q &:= \frac{\mathbf{y}_q}{M_{\mathcal{X}}} \end{aligned} \tag{4.4.7}$$

The normalization in (4.4.7) results in normalized training and validation data sets $\tilde{\mathcal{X}} = \{\tilde{\mathbf{x}_p} | p = 1, 2, \cdots, N_t\}$ and $\tilde{\mathcal{Y}} = \{\tilde{\mathbf{y}_q} | q = 1, 2, \cdots, N_v\}$ respectively.

4.4.3 Feed-forward Neural Network Architecture

The proposed deep learning framework employs an FNN architecture with multiple hidden layers that help to learn the relationship between set of IRs among multiple source-receiver pairs and position of the OI. The FNN architecture is highlighted by a dotted block shown in Fig. 4.6. The FNN takes an $N_b \times N_s \times N_r$ dimensional input vector of preprocessed features that are obtained using (4.4.4), (4.4.5) and (4.4.7). This is followed by N_l hidden layers and an output layer. Notice that the output from the previous layer is connected to the input of the current layer and there is no feedback path between the layers, hence the term feed-forward. The output layer contains the estimate of the OI position denoted by vector $\hat{\mathbf{z}}_o$ corresponding to the OI located at $\mathbf{z}_o$.

Each hidden layer contains a fully connected (FC) layer of neurons followed by rectified linear unit (ReLU) layer [156]. For instance, the ${L_j}^{\text{th}}$ FC layer contains N_{u_j} number of neurons and has connections to all neurons in the layer from the previous input. This is followed by a ReLU layer that performs a nonlinear operation on the output of neuron from the previous FC layer and is given as

$$a\left(w\right)=\max\left(0,w\right) \tag{4.4.8}$$

where $a\left(w\right)$ denotes the ReLU activation function and w denotes the output of neuron from previous FC layer. The ReLU activation function provides a greater convergence rate and is also computationally less expensive than sigmoid and tanh functions [157]. The final hidden layer L_{N_l} is then connected to the output layer. Since the coordinates of position have continuous domain i.e., $\{\mathbf{z}_o \in \Re^3 | \mathbf{z}_o \succcurlyeq \mathbf{0}\}$, in order to estimate position the output layer is implemented as a regression layer [96].

The data points in $\tilde{\mathcal{X}}$ with corresponding positions $\mathcal{Z}_{\mathcal{X}}$ are used to train the FNN. The trained network is then used to estimate positions of the data points with unknown positions. In practice, once the network is trained then a position estimate of the OI can be obtained by the following steps.

1. Obtain a set of IR measurements i.e., $\mathcal{G}$ in (4.3.5), corresponding to the OI with unknown position $\mathbf{z}_o$.
2. Represent $\mathcal{G}$ as feature-based vector $\mathbf{f}_o$ using (4.4.3).
3. Preprocess $\mathbf{f}_o$ using (4.4.6) to obtain preprocessed vector $\tilde{\mathbf{f}}_o$.
4. Use preprocessed vector $\tilde{\mathbf{f}}_o$ as an input to the trained network in order to estimate position $\hat{\mathbf{z}}_o$ of the OI.

4.4.4 Network Training

The objective of training the network is to minimize the error between the predicted and actual position with respect to the training data, i.e., $\mathcal{X}$ and $\mathcal{Z}_{\mathcal{X}}$, [96] and to avoid overfitting [158]. The efficacy of network training is characterized by computing the mean square error (MSE) between the predicted and actual positions of the OI. Define e_p as the MSE when $\mathbf{z}_o = \mathbf{z}_p \in \mathcal{Z}_{\mathcal{X}}$, that is

$$e_p = \frac{1}{N_z} \sum_{i=1}^{N_z} \left(\hat{z}_i^{(p)} - z_i^{(p)} \right)^2 \tag{4.4.9}$$

where $\hat{z}_i^{(p)}$ denotes the i^{th} predicted output and $z_i^{(p)}$ the i^{th} output coordinate corresponding to the OI at location $\mathbf{z}_p \in \mathcal{Z}_{\mathcal{X}}$, and N_z denotes the number of coordinates in the output layer.

The optimization algorithms in deep learning typically use a subset of training data points from the training data set $\mathcal{X}$ in order to train the network. The term *mini-batch* is used to refer to the subset of training data points. Define $\mathcal{X}_{B_i} \subset \mathcal{X}$ the i^{th} mini-batch that contains N_{B_i} training data points and f_{B_i} the corresponding mini-batch loss defined as [159]

$$f_{B_i} = \frac{1}{2N_{B_i}} \sum_{p=1}^{N_{B_i}} e_p. \tag{4.4.10}$$

The loss function for training data set is then defined as the average of all mini-batch losses and is given as

$$f_{\text{loss}} = \frac{1}{N_c} \sum_{i=1}^{N_c} f_{B_i} \tag{4.4.11}$$

where N_c are the total number of mini-batches.

It is important to note that only training data $\mathcal{X}$ and $\mathcal{Z}_{\mathcal{X}}$ are used to train the network, whereas the validation data $\mathcal{Y}$ are used only to evaluate the training performance and assess overfitting of the trained network. In order to compute validation loss, unlike training loss, the MSE is averaged over the complete data set $\mathcal{Y}$ instead of the mini-batches.

4.5 Simulation Results

4.5.1 Simulation Setup

The room depicted in Fig. 4.2 is modelled in Zemax using the parameters listed in Table 4.1. All of the N_s luminaries and N_r receivers are assumed to have identical parameters. The light sources and receivers are also assumed co-located on the ceiling. The data set of IR measurements are collected by placing the OI at randomly selected positions in the room. The region for selection of positions is chosen such that the areas covering furniture are excluded to avoid overlap of OI with the furniture. The total number of randomly selected positions in the room are $N_t + N_v = 1000$ to generate the data set of IR waveforms, where the duration of each IR waveform is $t_{\max} = 100$ ns. The number of data points in validation data set $\mathcal{Y}$ are set to $N_v = 200$. Define N_{t_i} as the number of training data points that are used to train the neural network and $N_{t_i} \leq N_t = 800$.

In relation to the measurement model of IRs in (4.3.4), the signal to noise ratio SNR is defined similarly to [131] for IR acquisition as

$$\mathrm{SNR} = \frac{\gamma^2 \left(\alpha P_s\right)^2}{\sigma^2} \tag{4.5.1}$$

Table 4.1: Parameter values for simulation

Room	- Dim. $(L \times W \times H)$, $5\,\mathrm{m} \times 5\,\mathrm{m} \times 3\,\mathrm{m}$ - Number of sources, $N_s = 9$ - Number of receivers, $N_r = 9$ - Reflection coefficient $(\rho_{\text{floor}}, \rho_{\text{wall}}, \rho_{\text{ceiling}})$, see Fig. 4.4, where ρ_{wall} uses plaster coating - Room temperature, 20 °C - Atmospheric pressure, 1 atm
Light Sources	- Total transmit power, $P_s = 1$ W - Lambertian index, $m = 1$ - Flux distribution vs wavelength, see Fig. 4.3
Receivers	- Responsivity, $\gamma = 1\mathrm{A/W}$ - half-angle FOV, $\Phi = 45\,\mathrm{deg}$ - Surface area, $A_r = 1\mathrm{cm}^2$ - Noise variance, $\sigma^2 = 10^{-16}\mathrm{A}^2$
Object of Interest	- Dim. $(L \times W \times H)$, $0.3\,\mathrm{m} \times 0.3\,\mathrm{m} \times 1.6\,\mathrm{m}$ - Reflection coefficient, $\rho_{\text{obj}} = 0.5$
Furniture	- Desk $(L \times W \times H)$, $0.5\,\mathrm{m} \times 1\,\mathrm{m} \times 0.85\,\mathrm{m}$ - Chair $(L \times W \times H)$, $0.43\,\mathrm{m} \times 0.41\,\mathrm{m} \times 0.95\,\mathrm{m}$ - Closet $(L \times W \times H)$, $1\,\mathrm{m} \times 1.6\,\mathrm{m} \times 1.85\,\mathrm{m}$ - Reflection coefficient, $(\rho_{\text{desk}}, \rho_{\text{chair}}, \rho_{\text{closet}})$, see pine wood coating in Fig. 4.4

where γ denotes responsitivity of the receiver measured in A/W, P_s (as mentioned in Sec. 4.3.1) is the maximum power emitted by the source in Watts, and $\alpha \in [0.1, 1]$ controls brightness of the light source i.e., 10% to 100%. The noise σ^2 is assumed independent and same at all the receivers. In order to evaluate VLP system performance, the root mean square error (RMSE) is calculated between the actual and estimated positions of the data points in validation data set $\mathcal{Y}$. The RMSE is given as

$$\text{RMSE} = \sqrt{\frac{1}{N_v} \left\| \mathbf{z}_q - \hat{\mathbf{z}}_q \right\|_2^2} \tag{4.5.2}$$

where $\mathbf{z}_q \in \mathcal{Z}_{\mathcal{Y}}$ and $\hat{\mathbf{z}}_q$ denotes the estimated positions of the OI and $N_v = |\mathcal{Y}|$.

The dimension of the input layer of the FNN is controlled by the number of time bins N_b as well as N_s and N_r. In the simulation setup in Fig. 4.2, $N_s = N_r = 9$. The parameters of the FNN are chosen after experimentation to ensure convergence and avoid overfitting. The number of hidden layers are $N_l = 8$ with each layer containing 500 neurons i.e., $N_{u_1} = N_{u_2} = \cdots = N_{u_8} = 500$. The size of the mini-batch to compute training loss is $N_B = 128$ elements. Since the data are collected for the OI of fixed height, the number of output variables in the output layer are set to $N_z = 2$. The deep learning toolbox in MATLAB [159] with *adam* optimizer is used to train the network.

4.5.2 Averaging IR Acquisitions for Noise Reduction

The noise in measured IRs can be reduced by accumulating IR acquisitions over N_{avg} number of times and then averaging the results [151]. As the measured IRs are typically on the order of 10's of ns, averaging IRs over several acquisitions still enable them to be obtained within fraction of a second e.g., the IR acquisition for one

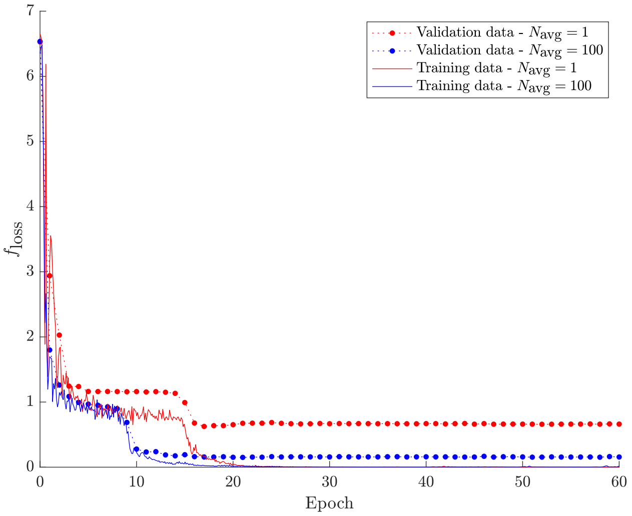


Figure 4.9: Comparison of loss function between training and validation data for averaging different number of IR acquisitions N_{avg}. The number of training and validation data points are $N_t = 800$ and $N_v = 200$ respectively. The time bin size is $t_b = 20$ ns ($N_b = 5$). The luminaires are set at full brightness ($\alpha = 1$).

source-receiver pair in case of $N_{\text{avg}} = 2$ and duration $t_{\text{max}} = 100$ ns can be obtained in 200 ns. In this simulations, $t_b = 20$ ns (or $N_b = 5$) and light sources are set at 100% brightness i.e., $\alpha = 1$. The number of training and validation data points are $N_t = 800$ and $N_v = 200$ respectively.

Fig. 4.9 shows the loss function for training and validation data sets against the increasing number of epochs, where one *epoch* is defined as a complete pass of the training data set through the FNN in Fig. 4.6 to update the weights and biases of the network. As inferred from the figure, the loss functions become stable after a small number of epochs and hence no overfitting [158] is observed in the curves. The figure shows two sets of curves that correspond to two values used for averaging IRs i.e., $N_{\text{avg}} = 1$ for no averaging and $N_{\text{avg}} = 100$ for averaging over 100 acquisitions. The curves show that performance is improved by averaging IRs over several acquisitions

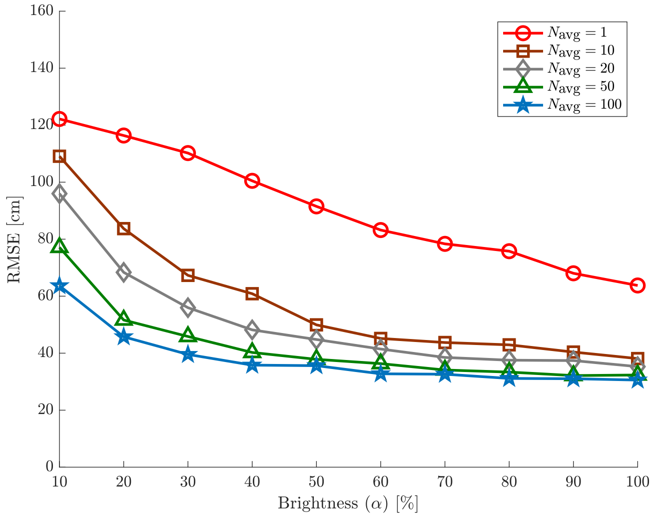


Figure 4.10: RMSE against the increasing brightness of light sources α. The number of training and validation data points are $N_t = 800$ and $N_v = 200$ respectively. The size of time bins is $t_b = 20$ ns ($N_b = 5$).

before the network is trained for position estimation at the expense of latency. The latency for each IR measurement $\mathcal{G}$ increases from 900 ns for $N_{\text{avg}} = 1$ to 90 μs for $N_{\text{avg}} = 100$.

It is important to note that both N_{avg} and α essentially change SNR, however, α is in control of the user in contrast to N_{avg}. Therefore, a small increase in latency due to large N_{avg} can considerably aid in mitigating the effects of noise.

4.5.3 RMSE vs Brightness of Light Sources

Fig. 4.10 shows RMSE plotted against the increasing brightness of light sources for different values of N_{avg}. In this simulation $t_b = 20$ ns (i.e., $N_b = 5$) and the number of training and validation points are $N_t = 800$ and $N_v = 200$ respectively.

The localization performance is sensitive to the brightness of the light sources.

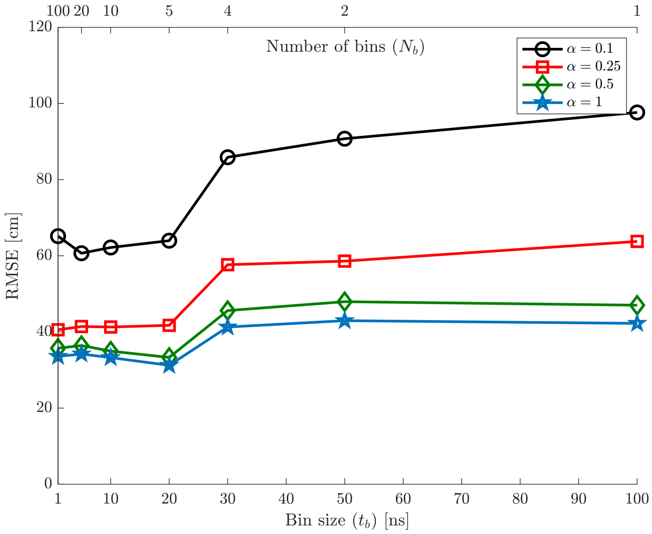


Figure 4.11: RMSE against the increasing size of time bins t_b or equivalently, the decreasing number of bins N_b for different brightness of light sources. The number of training and validation data points are $N_t = 800$ and $N_v = 200$ respectively. The number of acquisitions for averaging IRs are $N_{\text{avg}} = 100$.

The different curves in the figure correspond to different values of N_{avg}, which indicates that the performance is greatly improved when IRs are averaged over several acquisitions. For instance, in the case of $N_{\text{avg}} = 100$ the RMSE is around 60 cm at only a 10% brightness level of the luminaires.

4.5.4 RMSE vs Size of Time Bins for Feature Representation

The larger the time bin size t_b the smaller the value of N_b which reduces complexity of IR acquisition at the receivers. The performance here is evaluated by changing size of the time bins t_b when converting IRs into feature-based vectors. Fig. 4.11 shows the RMSE in localization of the OI against the increasing size of time bins. The number of training and validation data points are $N_t = 800$ and $N_v = 200$ respectively and $N_{\text{avg}} = 100$. The different curves shown in the figure correspond to different values

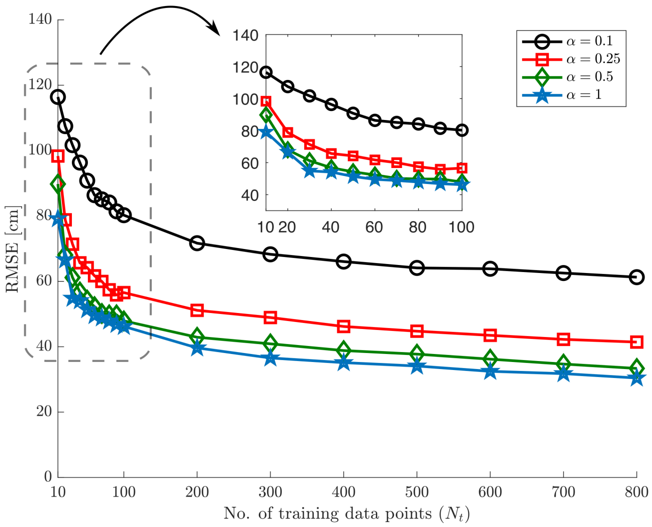


Figure 4.12: RMSE against the increasing number of training data points N_t for different values of α. The number of validation data points are $N_v = 200$. The number of runs for averaging IRs are $N_{\text{avg}} = 100$. The size of time bins is $t_b = 20$ ns ($N_b = 5$).

of brightness of light sources i.e., $\alpha = \{0.1, 0.25, 0.5, 1\}$.

The best performance is observed at full brightness of light sources where the variation in RMSE is around 10 cm i.e., 30 cm at $t_b = 20$ ns ($N_b = 5$) to 40 cm at $t_b = 100$ ns ($N_b = 1$). This indicates that the system can be operated at $t_b = 100$ ns ($N_b = 1$) with a small trade off in RMSE that results in reduction of the input dimensions by a factor of 5. The worst performance is observed at 10% brightness level of luminaires as expected, where variation in RMSE is around 40 cm i.e., 60 cm at $t_b = 5$ ns ($N_b = 20$) to near 1 m at $t_b = 100$ ns ($N_b = 1$). In this scenario, the time binning of IRs with small size is required to train the network in order to achieve the best performance.

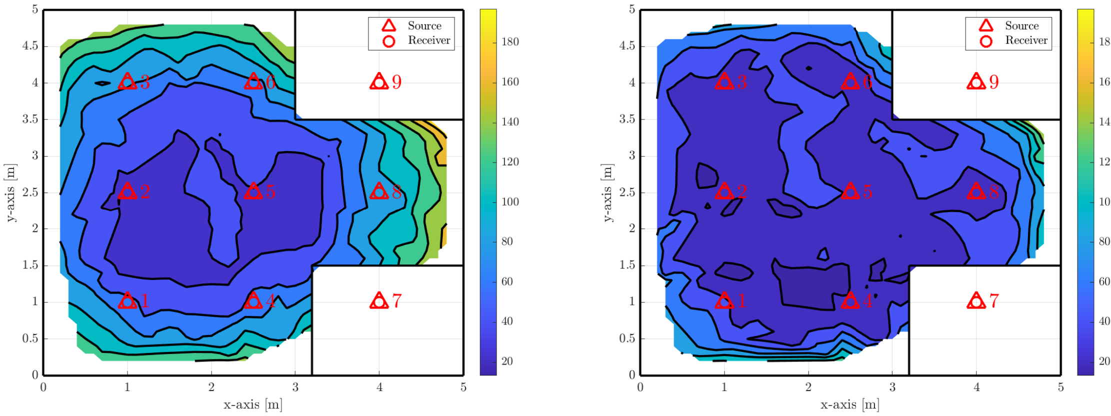
4.5.5 RMSE vs Number of Training Data points

In practice, the acquisition of training data points is a tedious and complex process. Fig. 4.12 shows the RMSE in OI localization (4.5.2) against the size of the training set N_t. The curves are obtained by averaging RMSE over 10 separate experimental runs of network training while using random combinations of the training data points in each run. The different curves in the figure correspond to different values of brightness of the light sources i.e., $\alpha = \{0.1, 0.25, 0.5, 1\}$. The number of IR acquisitions for averaging are $N_{\text{avg}} = 100$, $t_b = 20$ ns ($N_b = 5$) and the number of validation data points are set at $N_v = 200$.

Clearly, the performance is improved by increasing the number of training data points. A great improvement in accuracy can be observed at $N_t = 100$ after which the RMSE saturates. Notice, however, that the localization performance is still considerable even with a very small number of training data points i.e., the RMSE is 80 cm at $N_t = 10$. This is advantageous because it offers a huge reduction in the workload required to gather the data points and the system is still viable for many applications. If coarse localization of an OI is required then the deep learning framework can be trained using a limited number of training data points.

4.5.6 Spatial Distribution of RMSE in the Room

The localization performance depends on the position of OI in the room. In order to assess localization performance based on OI position, the spatial variation of RMSE in the room is plotted in Fig. 4.13 for different number of training data points, N_t. The RMSE plot shown in the figure is averaged over 10 separate experimental runs of network training while using random combinations of the training data in each run.



(a) (b)

Figure 4.13: Spatial variation of RMSE in the room for training data set of size: (a) $N_t = 10$ and (b) $N_t = 100$. The number of validation data points are $N_v = 200$ with same locations in both the figures. The number of IR acquisitions for averaging are $N_{\text{avg}} = 100$. The size of time bins of IRs is $t_b = 20$ ns ($N_b = 5$). The luminaires are set at full brightness i.e., $\alpha = 1$. The average RMSE in (a) and (b) corresponds to $\alpha = 1$ curve in Fig. 4.12 at $N_t = 10$ and $N_t = 100$ respectively.

The figure shows vacant areas in the top right and bottom corners, which are due to presence of furniture in those areas. These areas do not show any variation in RMSE since the OI is not placed in those regions while collecting the data.

The number of validation data points are $N_v = 200$ which are distributed uniformly in the room. The locations of validation data points are same in Figs. 4.13a and 4.13b. The number of acquisitions for averaging IRs are $N_{\text{avg}} = 100$, $t_b = 20$ ns ($N_b = 5$) and the luminaires are set at full brightness i.e., $\alpha = 1$.

The average RMSE values for $\alpha = 1$ at $N_t = 10$ and $N_t = 100$ are around 80 cm and 46 cm respectively, as apparent from Figs. 4.12 and 4.13. The comparison of figures shows that spatial performance is improved when a larger number of training data points are used to train the network. This can also be inferred that performance is relatively degraded when the OI is located near furniture of similar height e.g., closet in the top right corner. The performance also gets worse when the OI is located towards the edges of the room. This is due to the fact that reflections from the wall around same height as the OI are also accounted for in the IR. Moreover, at the edge of the room only a few receivers can sense the presence of OI in their vicinity due to their limited FOV.

4.6 Conclusions

A passive visible light positioning (VLP) system is proposed that employs a deep learning framework to estimate the position of an object of interest. The impulse responses (IRs) including multiple reflections among multiple source-receiver pairs affixed on the ceiling are constructed by using the ray trace data obtained from commercial optical system design software and realistic parameters. The presence of an OI inside the room creates small changes in IRs among the source-receiver pairs that can be related to the position of the OI. The deep learning framework learns the relationships between changes in IRs by the OI in order to estimate its position. It is important to note that although the proof-of-concept system proposed here considers a single OI in the room in order to train the deep learning framework, the approach and framework has interesting future extension to more general scenarios.

Simulation results demonstrate that the performance is greatly improved by using luminaires at full brightness as well as by averaging IRs over a large number of acquisitions. This averaging over a number of IR measurements increases SNR and leads to a tolerable increase in system latency due to the short time duration of the IR measurement in a room. The choice of large time bin size has an added advantage of reduced input dimensions when the luminaires are set at full brightness. The performance improvement can also be observed when the number of training data points used in training the network are increased. However, the performance at a small number of training data points is still sufficient for many applications i.e., the reported RMSE is around 80 cm at as low as 10 training data points. This has the advantage of a vast reduction in workload required to gather the data points since in practice, gathering the data points is a time consuming and difficult process. Furthermore, if coarse localization is required then the network can be trained by using a very small number of training data points. This passive VLP approach is an enabling localization approach for indoor IoT monitoring of patients in a hospital or home environment. Our future work considers experimental demonstration of this approach and location estimation of multiple OIs simultaneously in the room to make the system more widely applicable in indoor scenarios.

Chapter 5

Passive Indoor Visible Light-based Fall Detection using Neural Networks

IEEE Copyright Notice

If accepted, the copyright of the material in this chapter will be held by the IEEE.

The research work presented in this chapter is prepared for submission [90] to the IEEE and is shown in the following.

- **K. Majeed** and S. Hranilovic, "Passive Indoor Visible Light-based Fall Detection using Neural Networks," prepared for submission to the IEEE.

The contents of the paper are formatted according to the thesis requirements i.e., formatting of figures, equation numbers and settings, tables, etc.

In this chapter, a fall detection system is developed that classifies the orientation state of a user i.e., upright or prone in a room. This chapter employs same room model used in Chapter 4. Moreover, the process of impulse response (IR) acquisition is also similar to the previous chapters. The localization object (LO) i.e, a person is assumed present inside the room and considered to have one of the following states: upright or prone. The prone state of the LO is considered as a person lying on the floor along x or y axes.

The IR measurements corresponding to upright and prone states are labeled as class 1 and 2 respectively. The neural network takes in labeled data as input for training, where the labeled data is collected for the LO at uniformly distributed positions in the room in either states of the LO. In order to classify the state of the LO, the trained network takes a set of IR measurements as input and predicts the output class. The robustness of the proposed method is also evaluated by using a tilted state that is labeled with the same class number as the upright state. The results show that the trained network is able to differentiate the tilted state from the prone state even when the tilted state is not used to train the network. The reported accuracy is greater than 97% when the luminaires are set at full output power and the network is trained with only upright and prone states of the LO.

One of the strengths of machine learning algorithms is their ability to correctly predict the output when unseen measurements are fed to the trained network. The trained network should be robust to small changes in the indoor environment. For example, the tilted state considered in this chapter along four directions (north, east, south, and west) changes the IR measurements between the source-receiver pairs in comparison to the upright state. Though the tilted state is not used to train the network, the results show that the algorithm is correctly able to distinguish between the upright/tilted and prone states.

Secondly, the LO in prone state is considered lying along two directions only for the simplicity of data generation and simulations but in fact it can be oriented in any direction on the floor. A comparison of the IRs between upright and prone states of the LO show that the peaks corresponding to LO in the prone state occur much later in time with relatively smaller amplitudes as compared to the upright state and lie in the similar time range even when the LO is oriented along any direction on the floor. It is important to note that the path lengths of the traced rays increase considerably for the LO in prone state in contrast to the upright or tilted states.

5.1 Abstract

In this paper, a passive fall detection system based on visible light luminaires is proposed that uses neural networks to learn the state (upright or prone) of a target object (TO) e.g., a person. The proposed method measures the impulse response (IR) between different source-receiver luminaire pairs in a passive scenario, where the user does not hold a device or sensor. The IR measurements are collected in a realistically modeled room and neural networks are employed to learn the relationship between the IR measurements and the states of the TO at randomly selected positions in the room. The performance evaluation of the system shows that an accuracy of more than 97% is attainable by utilizing a large number of data samples and high brightness factor of the luminaires. The robustness of the proposed method is validated by using a tilted state which is labeled with same class as the upright state, however, the tilted state is not used to train the network. The correct prediction of the prone state is particularly critical in health care settings where emergency situations may arise from a fall.

5.2 Introduction

Visible light positioning (VLP) systems [46, 53] have attracted much attention recently due to the common availability of luminaires in the indoor environment. VLP techniques can be broadly classified into active and passive localization techniques based on the involvement of a user in the localization process.

In *active* localization, the user participates in the localization process by carrying a mobile device or a receiver and transmitting signals that are measured at the receiver

[46]. The energy in line-of-sight (LoS) components of the detected signals play an important role in positioning. However, a direct LoS is not necessarily required in *passive* localization scenarios [53, 87, 88] where the user does not hold any device, which is in contrast to more conventional active localization.

The detection of a human fall in a passive scenario can be considered as a complementary application of passive localization techniques [87, 88, 89] since it does not require additional hardware resources but rather employs signal measurements that are already available in the passive VLP methods. The authors in [160, 161] provide a comprehensive overview of the fall detection techniques. A majority of approaches require the use of devices worn by an individual to detect the fall. These devices include cameras, sensors attached to the user's body e.g., wrist, waist, ankle, etc., remote sensing receivers or a combination of aforementioned devices. Moreover, the privacy of the user is also at risk especially when using camera systems in an indoor area. The method proposed in [162] uses a purpose built motion depth sensor to detect human fall, where the data from the sensor are used to compute velocity of the object. The anomalous change in computed velocity is used to classify the fall. The methods in [163, 164] use cameras to obtain image or video of a person in order to detect the fall. The useful features required to classify the fall are extracted from the images captured by the camera. In [165], an infrared-based motion detector is used to detect a fall. The method in [166] employs an array of ultrasonic sensors that is used to perform gesture analysis of human posture in order to detect the fall.

In this paper, a fall detection system is proposed that uses existing luminaires in a room in order to classify state (i.e., upright or prone) of a target object (TO) (i.e., a person) in the indoor environment. The proposed infrastructure is similar to the passive scenario proposed in [87, 88, 89]. In order to classify the state of TO, impulse response (IR) measurements between source-receiver pairs are obtained for upright and prone states of the TO at uniformly distributed locations in the room. Neural networks are then applied to learn the relationship between the IRs and the states of the TO. The proposed method does not require the user to wear a sensor or hold a device in order to detect the fall as compared to many earlier methods [167, 160, 161]. Furthermore, the proposed method leverages existing lighting infrastructure, preserves the privacy of the user, and offers low complexity in predicting the fall due to the collection of IR measurements in a short period of time i.e, μs. It is important to note that, in order to detect the fall, the TO is assumed to be present inside the room. One of the important applications of fall detection techniques is to detect falls in healthcare settings since such events can result in serious injuries or even death in case of delayed emergency response.

The remainder of the paper is organized as follows. Sec. 5.3 describes modeling of a realistic room scenario. Sec. 5.4 shows performance evaluation of the proposed method. Finally, the paper is concluded in Sec. 5.5.

5.3 System Model

An example of passive scenario is shown in Fig. 5.1, where all three states (upright, tilted, and prone) of a user are shown with multi-order reflections of light rays between a source-receiver pair. The light rays w_1, w_2, w_3, and w_4 undergoing 2^{nd}-order, 1^{st}-order, 1^{st}-order, and 3^{rd}-order reflections respectively, between the source s_i and receiver r_j are also shown in the figure. This concept is realized by modeling a room in Zemax® Opticstudio [132] as described in Sec. 5.3.1.

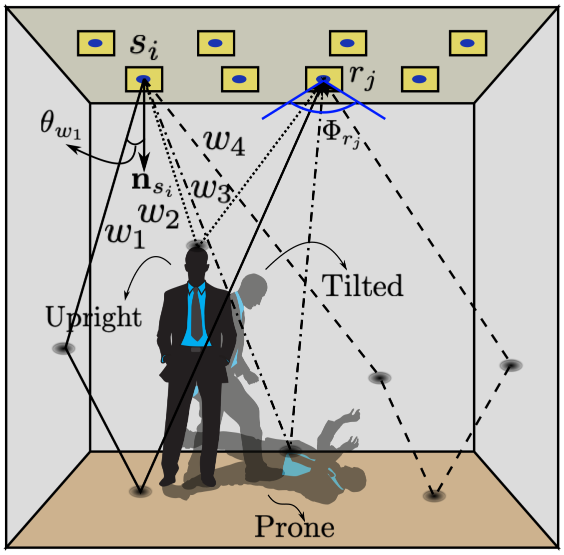


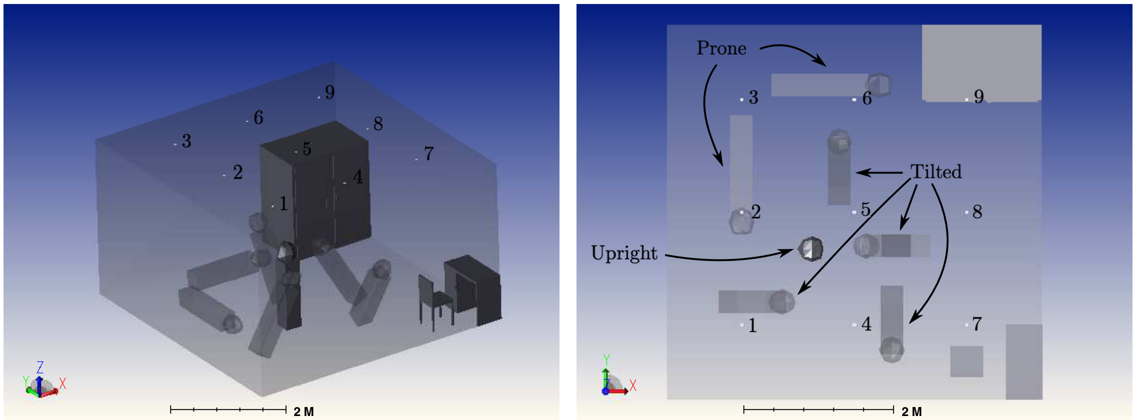
Figure 5.1: An example room scenario showing all three states (upright, tilted, and prone) of a user in the room. The rays w_1, w_2, w_3, and w_4 show 2nd-order, 1st-order, 1st-order, and 3rd-order reflections respectively. The silhouettes of a person are reproduced from [55].

5.3.1 Realistic Room Model

A realistic room environment similar to the one in [89] is modeled in Zemax® OpticStudio and is shown in Fig. 5.2 with the sources and receivers co-located on the ceiling. The figure shows an example of upright, tilted, and prone states of the TO at several random locations, where all possible cases of the tilted state (north, east, south, and west directions) and prone state (horizontal (x-axis) and vertical (y-axis) directions) are shown. It is important to note that the tilted state is not used in the training process, but rather used only to evaluate robustness of the proposed fall detection method.

The sources are assumed to have a Lambertian radiation pattern [149] given as

$$R\left(\varphi\right) = \frac{m+1}{2\pi}\xi P_{\max}\cos^{m}\left(\varphi\right) \tag{5.3.1}$$



(a) (b)

Figure 5.2: Zemax® model of the room containing co-located sources and receivers on the ceiling, furniture, and TO with upright, tilted, and prone states. The TO is located at (1.94 m, 2.01 m, 1.6 m) in upright position and room dimensions are 5m × 5 m × 3 m. The tilted (north, east, south, and west) and prone (horizontal and vertical) states at random locations are also shown in (a) 3D view and (b) top view of the room.

where m, ξ, and $P_{\max}$ are the Lambertian index, brightness factor, and total output power of the source respectively, and φ is the angle formed between the normal of the source and the ray emerging from the source. Moreover, the sources are also assumed to have variable flux distribution against wavelengths of visible light (i.e., photopic source in Zemax® [132]). The receivers have area A_r and field-of-view (FOV) Ψ. Furthermore, coating materials with measured wavelength-dependent characteristics are used to coat different surfaces in the room including furniture [95, 94, 93]. The CAD objects corresponding to the furniture are modeled using Blender [150].

Consider the IR $h_{s_i r_j}(\tau)$ between source s_i and receiver r_j that can be approximated as

$$h_{s_i r_j}(\tau) \approx \sum_{k=1}^{N_p} I_k \delta(\tau - \tau_k) \tag{5.3.2}$$

where N_p are the total number of rays detected at the receiver, I_k is the intensity of last

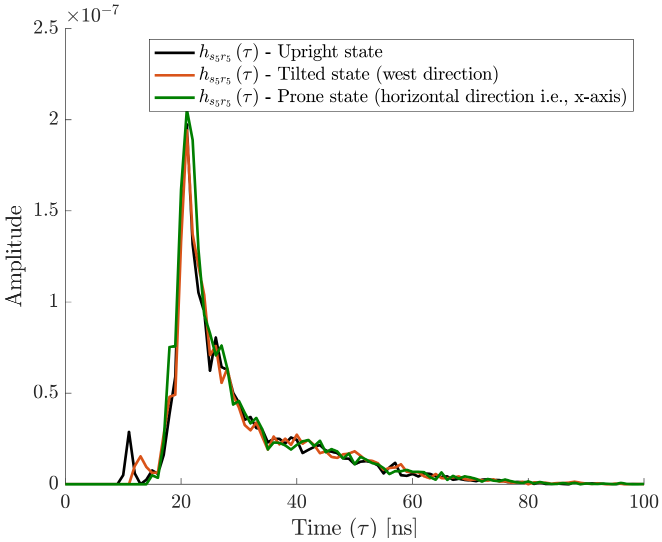


Figure 5.3: Impulse response $h_{s_5 r_5}(\tau)$ between source s_5 and receiver r_5 when the TO is located at (1.94 m, 2.01 m, 1.6 m) as illustrated in Fig. 5.2 for upright, tilted, and prone states (Note that the tilted and prone states are not shown in Fig. 5.2 for this TO location). The values of parameters used in simulation are listed in Table 5.1.

segment of the detected ray, and τ_k is the time corresponding to total path length of the detected ray. Examples of IR waveforms $h_{s_5 r_5}(\tau)$ corresponding to upright, tilted and prone states (west direction in case of the tilted state and horizontal direction i.e., along x-axis in case of the prone state) between source s_5 and receiver r_5 when the TO is located near center of the room (see Fig. 5.2) are shown in Fig. 5.3. The IRs can be well approximated as being time-limited, i.e. $\tau = [0, \tau_{\max}]$ and $\tau_{\max}$ can be chosen such that sufficient details of multi-order reflections are captured and is limited in practice by the storage available at the receiver.

The IR can be measured between a source-receiver pair by considering a single pair active at a time with measurement process repeated for all the pairs. The measured

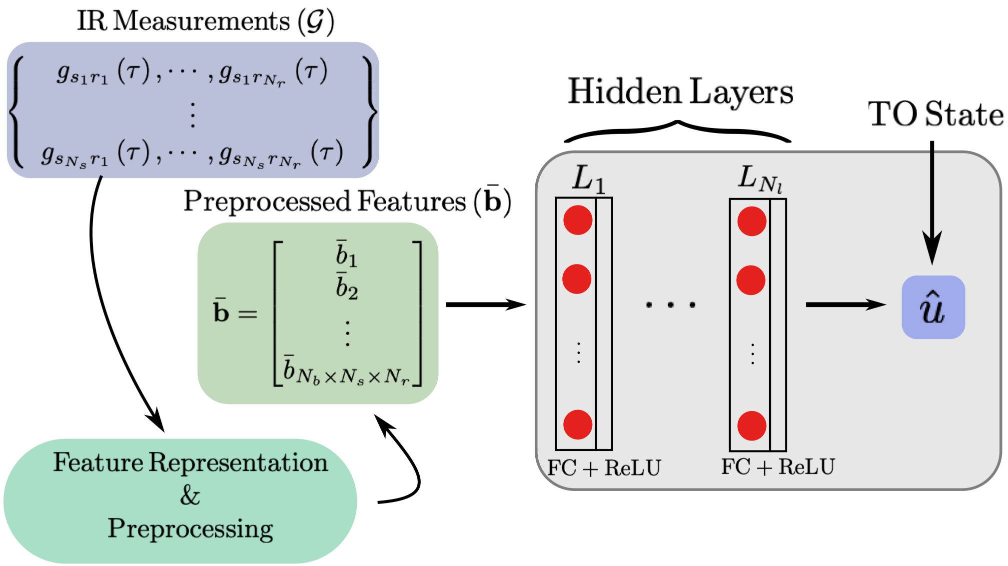


Figure 5.4: Block diagram showing a complete system architecture along with feed-forward neural networks for classifying output state of a TO.

IR $g_{s_ir_j}(\tau)$ between source s_i and receiver r_j can be modeled as

$$g_{s_ir_j}(\tau) = h_{s_ir_j}(\tau) + n_{s_ir_j}(\tau) \tag{5.3.3}$$

where $n_{s_ir_j}(\tau)$ is additive white Gaussian noise (AWGN) with zero mean and variance σ^2. Furthermore, the noise is considered to be independent amongst the receivers. The measured IRs between N_s sources and N_r receivers corresponding to a single state of the TO can be gathered in set $\mathcal{G}$ as

$$\mathcal{G} = \left\{ \begin{array}{c} g_{s_1r_1}(\tau), \cdots, g_{s_1r_{N_r}}(\tau) \\ \vdots \\ g_{s_{N_s}r_1}(\tau), \cdots, g_{s_{N_s}r_{N_r}}(\tau) \end{array} \right\} \tag{5.3.4}$$

5.3.2 Neural Network Classification

Neural networks are used to classify state of the TO i.e., upright or prone in the room. Fig. 5.4 shows a complete system model including the feed-forward neural networks (FNN) architecture [96] that is employed for classification of the state of the TO.

The IRs are measured in a similar fashion as shown in [89]. The measured IRs $g_{s_i r_j}(\tau) \in \mathcal{G}$ in (5.3.4) are divided into N_b time bins each where the time bins are defined as

$$\mathbb{T}_n = \begin{cases} [(n-1)\tau_b, n\tau_b), & \text{for } n = 1, 2, \cdots, N_b - 1 \\ [(N_b - 1)\tau_b, \tau_{\max}), & \text{for } n = N_b \end{cases} \tag{5.3.5}$$

The n^{th} feature of $g_{s_i r_j}(\tau)$ is defined as $b_{s_i r_j, n} = \int\limits_{\mathbb{T}_n} g_{s_i r_j}(\tau)\, d\tau$. The N_b features from all IRs are accumulated in an $N_b \times N_s \times N_r$ dimensional feature vector (data sample) as

$$\mathbf{b} = \begin{bmatrix} b_1 & b_2 & \cdots & b_{N_b \times N_s \times N_r} \end{bmatrix}^T \tag{5.3.6}$$

where the subscript of elements in $\mathbf{b}$ are simplified with correspondence to its dimensions.

Define $\mathcal{T}_{U,P}$ the training data set containing samples corresponding to upright and prone states only and $\mathcal{T}_{U,T,P}$ the training data set containing samples corresponding to all three states i.e., upright, tilted, and prone. The validation data set is defined as $\mathcal{V}_{U,T,P}$ and contains samples corresponding to all three states. In order to train the network, two independent cases of network training are considered that depend on the type of training data used, as described in Sec. 5.4. Define $\Gamma_{U,P}$ as the scenario when the network is trained using $\mathcal{T}_{U,P}$ and validated with $\mathcal{V}_{U,T,P}$. Similarly, $\Gamma_{U,T,P}$ refers to network training with $\mathcal{T}_{U,T,P}$ and validation using $\mathcal{V}_{U,T,P}$. It is important to note that the validation data set contains all three TO states i.e., $\mathcal{V}_{U,T,P}$ in both cases of network training.

The samples in $\mathcal{T}_{U,P}$, $\mathcal{T}_{U,T,P}$, and $\mathcal{V}_{U,T,P}$ are obtained using (5.3.4) and (5.3.6) and are first preprocessed before feeding them to the FNN architecture as indicated in Fig. 5.4. The vector elements in $\mathcal{T}_{U,P}$, $\mathcal{T}_{U,T,P}$, and $\mathcal{V}_{U,T,P}$ are simply represented by changing notations in (5.3.6). It is important to note that the samples in $\mathcal{T}_{U,P}$, $\mathcal{T}_{U,T,P}$, and $\mathcal{V}_{U,T,P}$ are obtained when the TO is located at randomly distributed locations in the room as described in Sec. 5.4. In order to preprocess, the samples in data sets $\mathcal{T}_{U,P}$ and $\mathcal{V}_{U,T,P}$ corresponding to the first case of network training $\Gamma_{U,P}$ are normalized as $\bar{\mathcal{T}}_{U,P} = \left\{\bar{\mathbf{t}}_{i(U,P)} = \frac{\mathbf{t}_{i(U,P)}}{M_{\mathcal{T}_{U,P}}} | i = 1, 2, \cdots, |\mathcal{T}_{U,P}|\right\}$ and $\bar{\mathcal{V}}_{U,T,P} = \left\{\bar{\mathbf{v}}_{j(U,T,P)} = \frac{\mathbf{v}_{j(U,T,P)}}{M_{\mathcal{T}_{U,P}}} | j = 1, 2, \cdots, |\mathcal{V}_{U,T,P}|\right\}$ respectively, where $\mathbf{t}_{i(U,P)}$ denotes the i^{th} sample vector in $\mathcal{T}_{U,P}$, $\bar{\mathbf{t}}_{i(U,P)}$ denotes the normalized i^{th} sample vector in $\bar{\mathcal{T}}_{U,P}$, $\mathbf{v}_{j(U,T,P)}$ denotes the j^{th} sample vector in $\mathcal{V}_{U,T,P}$, $\bar{\mathbf{v}}_{j(U,T,P)}$ denotes the normalized j^{th} sample vector in $\bar{\mathcal{V}}_{U,T,P}$, $M_{\mathcal{T}_{U,P}} = \max\limits_{1\leq l\leq N_b\times N_s\times N_r, 1\leq i\leq|\mathcal{T}_{U,P}|} t^{(i)}_{l(U,P)}$ is the maximum element in all vectors in $\mathcal{T}_{U,P}$, and the operator $|\cdot|$ denotes cardinality of a set. Similarly, the normalized data sets $\bar{\mathcal{T}}_{U,T,P}$ and $\bar{\mathcal{V}}_{U,T,P}$ are obtained for the second case of network training $\Gamma_{U,T,P}$.

The normalized training and validation data sets are fed to the FNN architecture through an input layer as illustrated in Fig. 5.4. This is then followed by N_l hidden layers and finally an output layer that predicts output state $\hat{u}$ of the TO. The parameter values used to model the FNN architecture are described in Sec. 5.4.

Table 5.1: Simulation Parameters

Model Room	- $L \times W \times H$ ($5\,\mathrm{m} \times 5\,\mathrm{m} \times 3\,\mathrm{m}$) - No. of sources and receivers ($N_s = N_r = 9$) - Reflection coefficient for floor, walls, and ceiling [93] - Room temperature and atmospheric pressure (20 °C and 1 atm)
Luminaires	- Maximum transmit power ($P_{\mathrm{max}} = 1$ W) - Lambertian index ($m = 1$) - Photopic source distribution [132]
Receivers	- Responsivity ($\alpha = 1$A/W) - half-angle FOV ($\Psi/2 = 45\,\mathrm{deg}$) - Surface area ($A_r = 1\mathrm{cm}^2$) - Noise variance ($\sigma^2 = 10^{-16}\mathrm{A}^2$)
Target Object (TO)	- ($L \times W \times H$) ($0.3\,\mathrm{m} \times 0.3\,\mathrm{m} \times 1.6\,\mathrm{m}$) - Reflection coefficient ($\rho_{\mathrm{obj}} = 0.5$) - Tilted state angle w.r.t. vertical axis (30 °) - Tilted state directions (north, east, south, west) - Prone state directions (horizontal (x-axis), vertical (y-axis))
Furniture	- Desk, $L \times W \times H$ ($0.5\,\mathrm{m} \times 1\,\mathrm{m} \times 0.85\,\mathrm{m}$) - Chair, $L \times W \times H$ ($0.43\,\mathrm{m} \times 0.41\,\mathrm{m} \times 0.95\,\mathrm{m}$) - Closet $L \times W \times H$ ($1\,\mathrm{m} \times 1.6\,\mathrm{m} \times 1.85\,\mathrm{m}$) - Pine wood coating [95, 94]

5.4 Numerical Results

5.4.1 Simulation Environment

The values of parameters used in defining the simulation environment for ray tracing in Zemax® are listed in Table 5.1, which corresponds to the room model shown in Fig. 5.2. The number of sources and receivers are $N_s = N_r = 9$ and are considered co-located on the ceiling. In order to collect the training ($\mathcal{T}_{U,P}$ or $\mathcal{T}_{U,T,P}$) and validation ($\mathcal{V}_{U,T,P}$) samples, the TO is located at randomly selected positions in the room for upright, tilted, and prone states and the IRs are measured between all source-receiver pairs for the TO in each state. The maximum time duration of IR is chosen as $t_{\max} = 100$ ns in order to capture sufficient details from multi-order reflections as evident from Fig. 5.3. The 100 ns time duration for IR acquisition corresponds to total path length around 30 m when a ray is traced from a source to receiver with multi-order reflections, which is sufficient given the dimensions of the room shown in Table 5.1.

The tilted state is simulated randomly along four directions i.e., north, east, south, and west and the prone state is modeled randomly along horizontal and vertical directions i.e., x and y axes. The total number of locations considered are 1000 with each position containing upright, tilted, and prone states, where the number of tilted and prone states are equally divided among their respective directions i.e., 250 each in case of tilted state and 500 each in case of prone state. The network is trained using two types of training data i.e., $\Gamma_{U,P}$ and $\Gamma_{U,T,P}$ separately in order to compare the performance between them as described in Sec. 5.3.2.

Define the state of the TO by u. In order to classify, the samples corresponding to the TO states are labeled as follows: upright state by $u = 1$ and prone state by $u = 2$. The tilted state is a transition between upright and prone states and it is considered as $u = 1$.

The noise model for signal-to-noise ratio (SNR) is considered similar to the one described in [131] and is given as

$$\mathrm{SNR} = \frac{\alpha^2 \left(\xi P_{\max}\right)^2}{\sigma^2} \tag{5.4.1}$$

where α is responsivity of the receiver measured in A/W, $P_{\max}$ is the total power transmitted by the source in W, $\xi \in [0.1, 1]$ controls brightness factor of the source, and the noise is considered independent among all the source-receiver pairs. Define N_a as the total number of IR acquisitions corresponding to a single source-receiver pair. The SNR can be improved by increasing N_a and then averaging the measurements [151, 89]. Notice that increasing N_a improves the SNR but results in increased latency for IR acquisitions.

The accuracy (%) of correctly predicting the TO state is calculated as

$$\text{Accuracy} = \frac{1}{|\mathcal{V}_{U,T,P}|} \sum_{j=1}^{|\mathcal{V}_{U,T,P}|} \mathbf{1}\left(\hat{u}_j = u_j\right) \tag{5.4.2}$$

where $\mathbf{1}(\cdot)$ is an indicator function that outputs 1 when the predicted and actual class values match and 0 vice versa and $|\mathcal{V}_{U,T,P}|$ are the total number of validation samples. The variables $\hat{u_j}$ and u_j denote the predicted and actual class values respectively for the j^{th} sample in validation data set $\mathcal{V}_{U,T,P}$. In order to avoid overfitting, the number of layers and neurons in each layer are chosen after experimentation, which are $N_l = 3$ with 500 neurons each and all the training samples are used as a single batch while

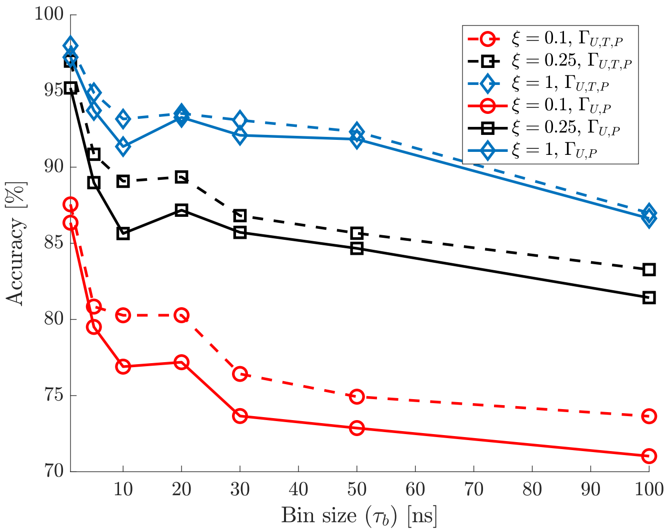


Figure 5.5: Accuracy against the increasing time bin size τ_b of the IRs (or equivalently decreasing no. of bins N_b) for different values of ξ of the luminaires. The number of training and validation samples for $\Gamma_{U,P}$ are $|\mathcal{T}_{U,P}| = 1600$ and $|\mathcal{V}_{U,T,P}| = 400$ respectively, and for $\Gamma_{U,T,P}$ are $|\mathcal{T}_{U,T,P}| = 2400$ and $|\mathcal{V}_{U,T,P}| = 600$ respectively.

training the network [96].

5.4.2 Accuracy Against the Increasing Time Bin Size τ_b

The performance is evaluated by changing time bin size τ_b of IRs in time binning process. The large τ_b reduces complexity of the analog-to-digital converter and storage requirements of the receiver at the detector. Fig. 5.5 shows accuracy against the increasing τ_b (or equivalently smaller number of bins N_b). The different curves in the figure correspond to different values of brightness factor ξ of the sources. The curves are obtained by averaging accuracy values over 30 separate experimental runs and using random training and validation samples in each run. The solid curves correspond to the first case of network training $\Gamma_{U,P}$ with $|\mathcal{T}_{U,P}| = 1600$ and $|\mathcal{V}_{U,T,P}| = 400$ and the dotted curves correspond to the second case of network training $\Gamma_{U,T,P}$

with $|\mathcal{T}_{U,T,P}| = 2400$ and $|\mathcal{V}_{U,T,P}| = 600$. It is important to note that the number of samples in validation set $\mathcal{V}_{U,T,P}$ are divided approximately evenly among the upright, tilted, and prone states. The number of IR acquisitions for a single source-receiver pair for averaging as defined earlier are $N_a = 150$. This results in increased system latency for collecting measurements in (5.3.4) i.e., 900 ns for $N_a = 1$ to 135 μs for $N_a = 150$ in case of IR with length $t_{\max} = 100$ ns and $N_s = 9$ sources since turning one source on at a time enables IR acquisitions at all the receivers. However, IR measurements between all source-receiver pairs can still be obtained within fraction of a second.

The performance improves by increasing brightness factor ξ of the luminaires as expected. This can be inferred from the figure that the improvement in accuracy is modest with the inclusion of the tilted state in the training, which shows robustness of the proposed approach. However, this improvement is at the cost of increased size of training set i.e., $|\mathcal{T}_{U,T,P}| = 2400$.

5.4.3 Accuracy Against the Increasing Brightness Factor ξ

The performance is also evaluated with respect to the increasing brightness factor ξ of the luminaires as shown in Fig. 5.6. The different curves in the figure correspond to different values of N_a used for averaging multiple IR acquisitions. The curves are obtained by averaging accuracy values over 5 independent experimental runs using random training and validation samples in each run for two separate cases of network training i.e., $\Gamma_{U,P}$ and $\Gamma_{U,T,P}$. The number of training and validation samples for both the training cases are same as used in Fig. 5.5. The time bin size used for binning IRs is $\tau_b = 1$ ns (or $N_b = 100$), which is highest bandwidth sampling as shown in

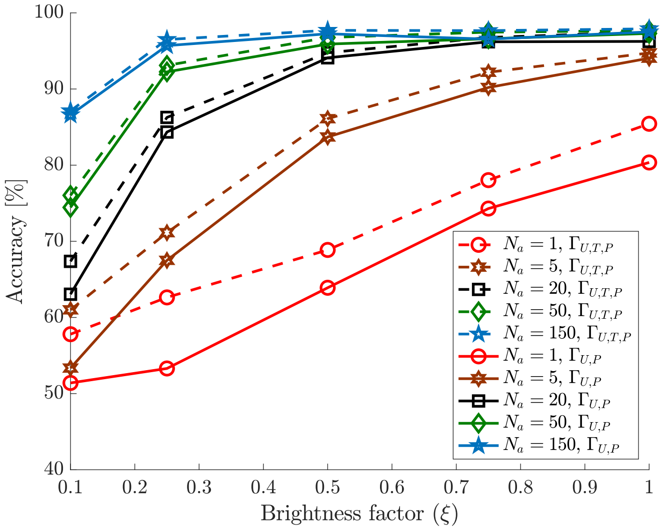
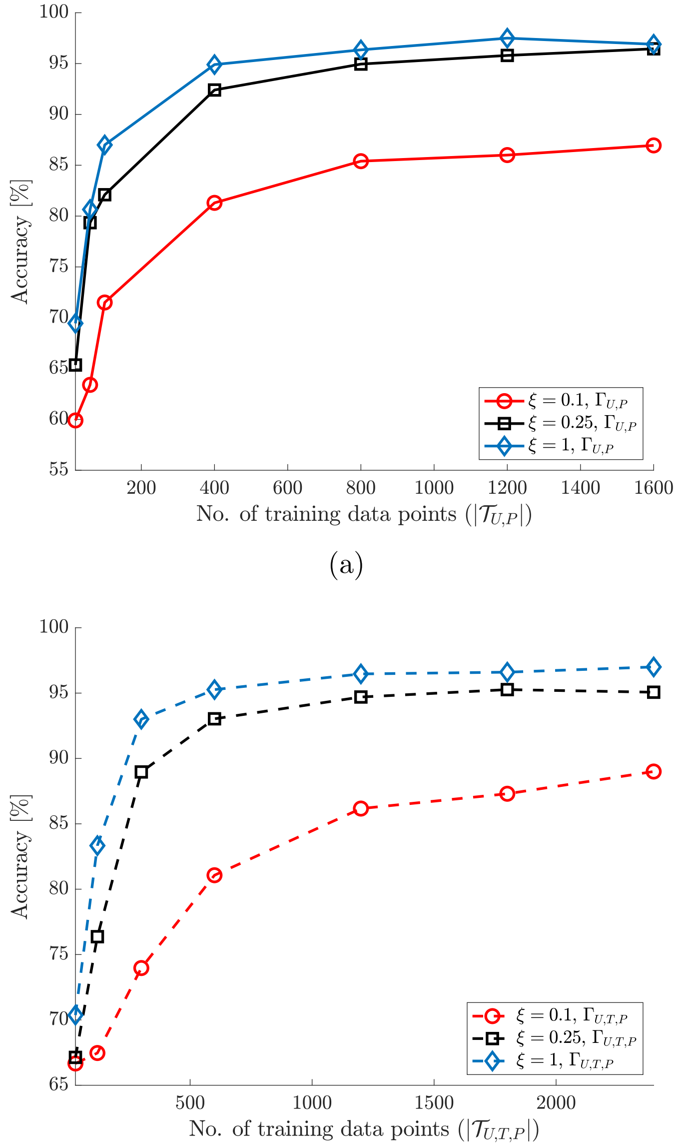


Figure 5.6: Accuracy against the increasing brightness factor ξ of the luminaires for different averaging values N_a of IR acquisition. The time bin size is $\tau_b = 1$ ns. The number of training and validation samples for $\Gamma_{U,P}$ are $|\mathcal{T}_{U,P}| = 1600$ and $|\mathcal{V}_{U,T,P}| = 400$ respectively, and for $\Gamma_{U,T,P}$ are $|\mathcal{T}_{U,T,P}| = 2400$ and $|\mathcal{V}_{U,T,P}| = 600$ respectively.

Fig. 5.5. The comparison of the curves show that performance of the system can be greatly improved by using IRs averaged over large N_a. The performance also improves by increasing brightness of the luminaires. Furthermore, the addition of a tilted state in training (dotted curves) i.e., $\Gamma_{U,T,P}$ shows comparable performance to using only upright and prone states (solid curves) i.e., $\Gamma_{U,P}$, especially for $N_a = 150$. This demonstrates that the system is robust to distinguish the tilted state from the prone state despite the fact of excluding tilted state in training the network.

5.4.4 Accuracy Against the Increasing Number of Training Samples

Fig. 5.7 shows accuracy against the increasing number of training samples for different brightness factor ξ of the luminaries. The time bin size for binning process of IRs is $\tau_b = 1$ ns ($N_b = 100$). The number of IR acquisitions for averaging are $N_a = 150$. The accuracy curves are averaged over 5 separate experimental runs. Fig. 5.7a shows network training $\Gamma_{U,P}$ with $|\mathcal{V}_{U,T,P}| = 400$ samples and Fig. 5.7b shows $\Gamma_{U,T,P}$ with $|\mathcal{V}_{U,T,P}| = 600$ samples. The training and validation samples are chosen randomly during each experimental run and the number of elements in $\mathcal{V}_{U,T,P}$ are divided approximately equally among the three states. As inferred form the figure, high accuracy can be achieved using a large training set with luminaires set at full brightness. The comparison of Figs. 5.7a and 5.7b shows that the performance of network training $\Gamma_{U,P}$ is comparable to $\Gamma_{U,T,P}$. Secondly, the performance saturates after $|\mathcal{T}_{U,P}| = |\mathcal{T}_{U,T,P}| = 800$ samples in both the figures. This indicates that the network can be trained with less number of samples instead of using the complete training data set.



(a)

(b)

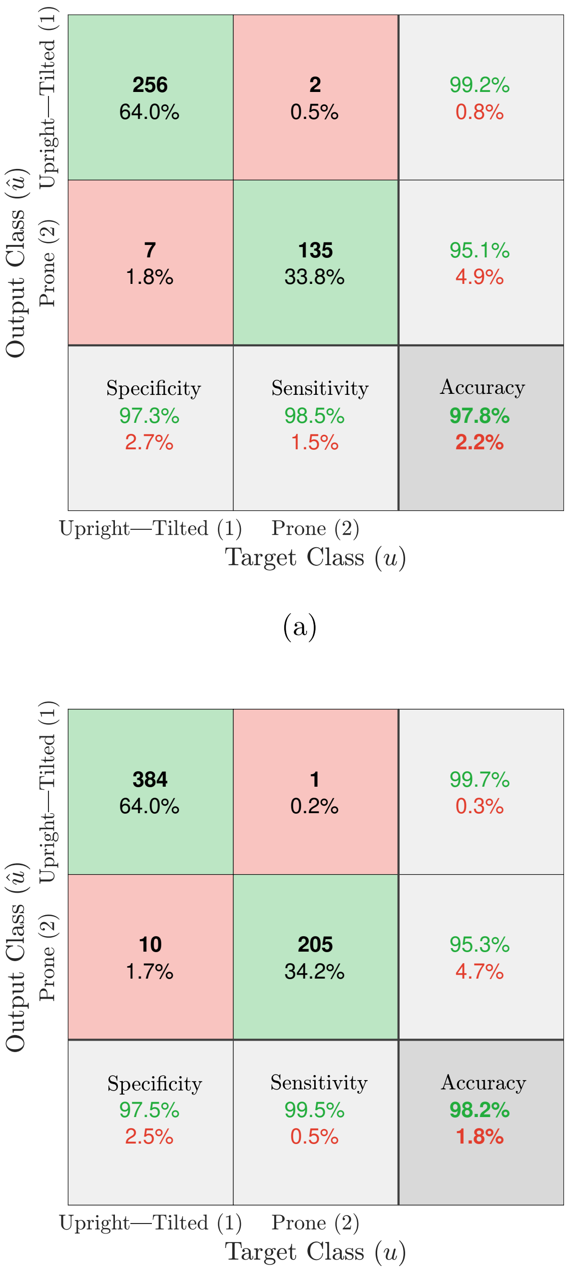
Figure 5.7: Accuracy against the increasing number of training samples for different brightness factor (ξ) of the luminaires. The time bin size is $\tau_b = 1$ ns and number of IR acquisitions for averaging are $N_a = 150$. (a) $\Gamma_{U,P}$ with $|\mathcal{V}_{U,T,P}| = 400$ samples and (b) $\Gamma_{U,T,P}$ with $|\mathcal{V}_{U,T,P}| = 600$ samples.

5.4.5 Confusion Matrix

Fig. 5.8 shows confusion matrix for one experimental run of network training using training and validation data which are uniformly distributed in the room. Fig. 5.8a uses $|\mathcal{T}_{U,P}| = 1600$ and $|\mathcal{V}_{U,T,P}| = 400$ samples while Fig. 5.8b uses $|\mathcal{T}_{U,T,P}| = 2400$ and $|\mathcal{V}_{U,T,P}| = 600$ samples. The figure depicts state prediction accuracy for both upright (or tilted) and prone states of the TO. The time bin size is $\tau_b = 1$ ns ($N_b = 100$) and number of IR acquisitions for averaging are $N_a = 200$. The brightness factor of the luminaires is set at $\xi = 1$. The accuracy of more than 97% can be observed in Fig. 5.8a when only upright and prone states are used in training the network i.e., $\Gamma_{U,P}$. On the other hand, the accuracy greater than 98% can be observed in Fig. 5.8b when all three states are used in network training i.e., $\Gamma_{U,T,P}$. The comparison of Figs. 5.8a and 5.8b shows that the system is robust in predicting the state of TO even when it is tilted.

The correct prediction of prone state is of utmost importance in fall detection techniques. Apart from accuracy values, sensitivity and specificity measures are typically used in order to assess the performance of fall detection systems. Sensitivity measure reflects the correct prediction of true positives i.e., it shows how well the system performs in correctly predicting the fall and specificity measure shows correct prediction of true negatives i.e., robustness of the system in rejecting incorrectly predicted falls [160]. The sensitivity and specificity values can be calculated as

$$\begin{aligned} \text{Sensitivity} &= \frac{N_{TP}}{N_{TP} + N_{FN}} \\ \text{Specificity} &= \frac{N_{TN}}{N_{TN} + N_{FP}} \end{aligned} \tag{5.4.3}$$



Output Class ($\hat{u}$) \ Target Class (u)	Upright—Tilted (1)	Prone (2)	
Upright—Tilted (1)	**256** 64.0%	**2** 0.5%	99.2% 0.8%
Prone (2)	**7** 1.8%	**135** 33.8%	95.1% 4.9%
	Specificity 97.3% 2.7%	Sensitivity 98.5% 1.5%	Accuracy **97.8%** **2.2%**

(a)

Output Class ($\hat{u}$) \ Target Class (u)	Upright—Tilted (1)	Prone (2)	
Upright—Tilted (1)	**384** 64.0%	**1** 0.2%	99.7% 0.3%
Prone (2)	**10** 1.7%	**205** 34.2%	95.3% 4.7%
	Specificity 97.5% 2.5%	Sensitivity 99.5% 0.5%	Accuracy **98.2%** **1.8%**

(b)

Figure 5.8: Confusion matrix showing prediction accuracy for upright/tilted ($u = 1$) and prone ($u = 2$) states. The time bin size is $\tau_b = 1$ ns ($N_b = 100$) and number of IR acquisitions for averaging are $N_a = 200$. The brightness factor of the luminaires is $\xi = 1$. Network training (a) $\Gamma_{U,P}$ with $|\mathcal{T}_{U,P}| = 1600$ and $|\mathcal{V}_{U,T,P}| = 400$ and (b) $\Gamma_{U,T,P}$ with $|\mathcal{T}_{U,T,P}| = 2400$ and $|\mathcal{V}_{U,T,P}| = 600$.

Table 5.2: Sensitivity and Specificity values for experimental run (Fig. 5.8)

	$\Gamma_{U,P}$	$\Gamma_{U,T,P}$
Sensitivity	98.5%	99.5%
Specificity	97.3%	97.5%
Accuracy	97.8%	98.2%

where N_{TP}, N_{TN}, N_{FP}, and N_{FN} represent number of true positives, true negatives, false positives, and false negatives respectively. Consider the prone state ($u = 2$) as true positive and the upright (or tilted) state as true negative ($u = 1$). The sensitivity, specificity, and accuracy values corresponding to $\Gamma_{U,P}$ and $\Gamma_{U,T,P}$ for the experimental run in Fig. 5.8 are shown in Table 5.2. The comparison of the performance measure values between $\Gamma_{U,P}$ and $\Gamma_{U,T,P}$ shows robustness of the fall detection system in predicting the prone state of the TO even when the tilted state is excluded in network training i.e., $\Gamma_{U,P}$.

The proposed method achieves accuracy greater than 97% and relieves the user from carrying a device or wearing a sensor, which is in contrast to the most fall detection methods surveyed in [168, 160, 161] that require wearable sensors, camera-based sensors, ambient sensors, or combination of them in order to detect the falls. The ambient sensors-based study in [161] describes methods that sense changes in the indoor environment e.g., using radio frequency signals i.e., WiFi [169]. The method in [169] uses channel state information of WiFi signals in order to detect the fall and the reported accuracy is around 93%. The reported accuracy in [170] is greater than 95%, however, the proposed method uses geophone sensors in the indoor area in order to detect floor vibrations that occur due to the fall. The sensitivity and specificity of the fall detection method in [168] is 96%, however, the method requires the user to wear accelerometer sensor.

5.5 Conclusions

A fall detection system using visible light in a realistically modeled indoor environment is proposed that uses neural networks in order to predict output state of a target object (TO). A set of impulse response (IR) measurements between different source-receiver pairs in the room are obtained that correspond to upright or prone states of the TO. The employed neural network architecture learns the relationship between a set of IR measurements and the TO states. The proposed method does not require any sensor tags attached to the user and predicts the fall assuming that the user is present inside the room.

The robustness of the proposed method is also evaluated by excluding tilted state in network training. The accuracy of more than 97% is reported when large number of training samples are used with luminaires set at full brightness. Our future work focuses on using dynamic falling motion models of a person in the indoor area and collecting series of measurements in order to track their state.

Chapter 6

Conclusions and Future Work

In this chapter, the contributions of the proposed research work are summarized. Furthermore, possible future directions are also discussed in relation to the visible light-based passive indoor localization paradigm.

6.1 Conclusions

In this thesis, a proof-of-concept of passive indoor localization system is proposed that uses visible light as the underlying technology. The passive localization scenario, in contrast to active localization, does not involve a user directly in the localization process but rather sense changes in the indoor environment in order to estimate position of the user. The user neither carries a device with them nor do they have any sensor tags attached to their body. A network of luminaires and receivers are installed in the room that can be controlled from a central back end server. An example of controlling such a network can also be through Internet-of-things devices that allows the control of devices over the Internet. Moreover, the smart buildings concept encapsulates control of luminaires through power-over-Ethernet (PoE), where the network is able to provide communication and power to the luminaires through the same channel. In order to acquire signals, the impulse response (IR) measurements are obtained between all source-receiver pairs in the room. The presence of a localization object (LO) inside the room produces changes in IRs between certain source-receiver pairs depending on its position. The work presented in this thesis can be divided into two parts based on the considered room model i.e., single-bounce model and realistic room model.

In the single-bounce model, an empty room is considered with first-order reflections only. The IR due to LO is obtained by computing difference between the IRs with and without the presence of LO in the room. An exponential integrating-sphere model is fitted to the IR differences and a fingerprinting method is used to estimate the position of the LO. Furthermore, the Cramér-Rao lower bound is derived on the positioning error and the tightness of the bound is evaluated by using a maximum likelihood estimator.

In a realistic scenario, the room is modeled with reasonable parameters in an optical system design software where multi-order reflections are also included in the measured IRs. The sources have wavelength-dependent characteristics. The optical properties of surfaces in the room are modelled by measured wavelength-dependent reflectivity coatings on all surfaces of the room i.e., walls, floor, ceiling, and furniture. The furniture is modeled using a 3D graphics design software and exported as computer-aided design objects that are placed inside the modeled room.

The presence of the LO inside the room produces changes in measured IRs between the source-receiver pairs. Changes in the IRs correspond to the LO present at a certain position in the room. A deep learning framework is employed to learn these changes in order to estimate the position of the LO at an unknown location. Secondly, a passive fall detection system is proposed that classifies state of the LO i.e., upright or prone. Neural networks are used to learn the state of LO from the IR acquisitions. The trained network then classifies state of the LO from the IR measurements.

One of the proposed passive positioning solutions employs deep learning for position estimation in a realistic indoor scenario. It is likely that the performance between classical machine learning techniques is comparable to deep learning techniques given the simple room layout. However, this may not be the case in more realistic indoor scenarios with enriched furnishings or considerably larger extent. The use of deep learning in these scenarios offers much larger benefits since they are able to learn minute details of the indoor environment and outperform classical machine learning algorithms.

The Passive visible light positioning (VLP) approach protects privacy of the user while simultaneously providing positioning service. The applications of the proposed passive VLP system include monitoring of patients in a hospital environment or at home, surveillance of people in private indoor areas, detecting workers in high-risk indoor environments, etc.

Passive positioning based on visible light relies on LED luminaires in order to estimate the position of the user. Therefore, the system cannot work in the absence of light or at night, however, it can be modified to include IR sources and sensors because the wave propagation characteristics in IR spectrum are analogous to visible light spectrum. Therefore, the algorithms developed for visible light positioning can be used analogously for IR positioning.

In terms of complexity, fingerprinting approaches are computationally complex because they require exhaustive search in a stored map in order to estimate the position of the user. The deep learning approach has larger computational complexity only in the network training phase. However, the network training needs to be done only once for a specific room. Once the network is trained then the computational complexity for position estimation involves only multiplications and additions that is a function of number of inputs, weights and biases in the network.

6.2 Future Work

The passive visible light positioning approach is proposed for both single-bounce and realistic indoor environments that does not require a user to carry any device or sensor tags. However, there are some possible research directions that can be explored potentially in the future work.

6.2.1 User Tracking

The passive positioning approach considers position estimation of a user in an indoor environment at a stationary location. However, the proposed work can be extended to include real-time signal measurements in order to track movement of the user in the indoor area using tracking algorithms e.g., extended Kalman filter (EKF). The approach can also take into account motion models of a person walking in the area in order to improve efficiency of tracking.

Another direction could be differentiating between a stationary and moving users by collecting a series of IR measurements. This can be achieved by tracking peaks in the IR measurements corresponding to a user in order to detect their motion.

6.2.2 Position Estimation of Multiple Users

The proposed research work considers a single user in an indoor environment whose position is estimated. The work can be extended to measure IRs when multiple users are present in the room. The users present in respective localities of the room cause changes in IRs in the source-receive pairs in their vicinity. Positioning algorithms can be developed that relate changes in the IR measurements with locations of the users in order to estimate their positions. An advanced deep learning framework e.g., convolutional neural networks (CNNs) can be employed to learn about the number of peaks in IRs corresponding to the number of users present in a realistically modeled indoor environment since CNNs exploit spatial information in the training data.

Secondly, the IR differences obtained in Chapters 2 and 3 have their negative parts zero-clipped. The energy in the negative part depends on the location of a user as well as the size of their shadow formed in the room. The negative part in the IR differences can also be utilized and appropriately modeled in order to relate it uniquely to a single user.

6.2.3 User count detection

Some indoor applications require calculation of user density in an indoor area in order to provide communication services e.g., video streaming, web browsing, etc. The density can be calculated by implementing a count estimator system that is able to count the number of users simultaneously present in the indoor area. The IR measurements can be obtained that show changes in them corresponding to multiple users similar to the one described in Sec. 6.2.2. The location of the user near a certain source-receiver pair produces prominent peaks in first-order bounces as compared to the higher order bounces. These peaks can be detected in order to count the number of users in the area.

6.2.4 Scalability and Viability of Passive Indoor Positioning System in Multiple Indoor Environments

The proposed passive localization method considers a room with a fixed layout. However, one of the challenges in implementation of indoor localization systems is that they can be easily deployed in new indoor areas. Although machine learning algorithms help to learn the layout of a particular room, however, in order to deploy systems in new indoor areas, a transfer learning framework can be utilized. The transfer learning approach learns positioning scenario in a room with specific layout and the trained network can be used as a prior to predict outputs in a room with different layout. Nevertheless, the transfer learning approach can help to implement the indoor passive positioning system in different indoor environments without any hassle, hectic planning or resource requirements.

Secondly, the proposed fall detection system considers a static state of a person at a given time instant i.e., upright or prone. The system can be improved to include measurements from dynamic model of a falling person. For example, the measurements from two different states can be concatenated while training the network. The concatenation of different states can provide more insights into observing the fall through user motion as investigated in camera-based systems.

6.2.5 Experimental Validation

The proposed work presents a proof-of-concept of passive indoor localization scenario using single-bounce and realistic room models. Though the realistic room model considers the modeling of a room in an optical system design software that considers the environment as realistic as possible, however, an experimental setup can be used to verify integrity of the proposed approach. The real luminaires and photodetectors (PDs) can be used to implement such a system. The IR measurements can be obtained by controlling luminaires and PDs through a central server.

The major challenge in the deployment of passive localization systems is the design or selection of luminaires and PDs. The devices should be selected such that their bandwidth allows to capture enough details corresponding to the presence of user in the room. Secondly, each source should be synchronized with all the receivers through some backbone network. However, no precise synchronization is required between the sources.

Appendix A

Supplementary Material – Performance Bounds on Passive Indoor Positioning using Visible Light

This appendix is related to the research work described in Chapter 3 and was submitted as supplementary material for the following journal.

- [88] **K. Majeed** and S. Hranilovic, "Performance Bounds on Passive Indoor Positioning Using Visible Light," in *Journal of Lightwave Technology*, vol. 38, no. 8, pp. 2190-2200, 15 April 15, 2020, doi: 10.1109/JLT.2020.2966365.

This appendix contains detailed derivation of the entries in Fisher information matrix that are required to calculate analytical expression of the Cramér-Rao lower bound. The appendix is reproduced in the following with minor modifications in order to follow the thesis format.

A.1 Introduction

This supplement to our submission [88] contains additional details on the partial derivatives derived in the computation of the Cramér-Rao lower bound (CRLB) for indoor passive positioning systems, shown in Section 3.4. The same notation is employed in this supplement as is defined in the main submission. The definitions of important variables, model and its parameters are stated here again for easy reference.

The positions of source s_i and receiver r_j are represented by vectors $\mathbf{x}_i^{(s)} = [x_{s_i}\; y_{s_i}\; z_{s_i}]^T$ and $\mathbf{x}_j^{(r)} = \left[x_{r_j}\; y_{r_j}\; z_{r_j}\right]^T$ respectively. Let $\mathbf{x}_o = [x_o\; y_o\; z_o]^T$ represent center position of top surface of the object. The OIR model from Section 3.3.2 is given by

$$\tilde{h}_{s_i r_j}(t, \mathbf{x}_o) = \eta_{s_i r_j}(\mathbf{x}_o)\, e^{-\frac{\left(t - t_{s_i r_j}(\mathbf{x}_o)\right)}{\tau_{s_i r_j}(\mathbf{x}_o)}}\, u_{\sigma_a,\varepsilon}\left(t - t_{s_i r_j}(\mathbf{x}_o)\right) \tag{A.1.1}$$

The amplitude $\eta_{s_i r_j}(\mathbf{x}_o)$, time delay $t_{s_i r_j}(\mathbf{x}_o)$ from center of object's top surface, and decay rate $\tau_{s_i r_j}(\mathbf{x}_o)$ are defined as

$$\eta_{s_i r_j}(\mathbf{x}_o) = (\beta P_s)\frac{(m+1)\,\rho_o A_{r_j} A_{o,\text{eff}}}{2\pi} \times \frac{(z_{s_i} - z_o)^{m+1}\left(z_{r_j} - z_o\right)^2}{\left\|\mathbf{x}_o - \mathbf{x}_i^{(s)}\right\|_2^{m+3}\left\|\mathbf{x}_j^{(r)} - \mathbf{x}_o\right\|_2^4}, \tag{A.1.2}$$

$$t_{s_i r_j}(\mathbf{x}_o) = \frac{\left\|\mathbf{x}_o - \mathbf{x}_i^{(s)}\right\|_2}{c} + \frac{\left\|\mathbf{x}_j^{(r)} - \mathbf{x}_o\right\|_2}{c}, \tag{A.1.3}$$

and

$$\tau_{s_ir_j}(\mathbf{x}_o) = \frac{1}{2} \times \Big(\big(t^{(a_1)}_{s_ir_j}(\mathbf{x}_o) - t_{s_ir_j}(\mathbf{x}_o)\big)^2 + \big(t^{(a_2)}_{s_ir_j}(\mathbf{x}_o) - t_{s_ir_j}(\mathbf{x}_o)\big)^2 + \big(t^{(a_3)}_{s_ir_j}(\mathbf{x}_o) - t_{s_ir_j}(\mathbf{x}_o)\big)^2 + \big(t^{(a_4)}_{s_ir_j}(\mathbf{x}_o) - t_{s_ir_j}(\mathbf{x}_o)\big)^2 \Big)^{\frac{1}{2}} \quad \text{(A.1.4)}$$

respectively.

A.2 Definition of delays for $\tau_{s_ir_j}(\mathbf{x}_o)$

In order to compute $\tau_{s_ir_j}(\mathbf{x}_o)$, the time delays from each of the corners of the top surface of the object, $\mathcal{S}_o$, are needed, as shown in Figure 2 of the paper. The time delays from corners a_1, a_2, a_3, and a_4 of $\mathcal{S}_o$ are given explicitly in (A.2.1), (A.2.2), (A.2.3), and (A.2.4) respectively.

$$t^{(a_1)}_{s_ir_j}(\mathbf{x}_o) = \frac{\sqrt{\left((x_o - \Delta x) - x_{s_i}\right)^2 + \left((y_o + \Delta y) - y_{s_i}\right)^2 + \left(z_o - z_{s_i}\right)^2}}{c} + \frac{\sqrt{\left(x_{r_j} - (x_o - \Delta x)\right)^2 + \left(y_{r_j} - (y_o + \Delta y)\right)^2 + \left(z_{r_j} - z_o\right)^2}}{c} \quad \text{(A.2.1)}$$

$$t^{(a_2)}_{s_ir_j}(\mathbf{x}_o) = \frac{\sqrt{\left((x_o + \Delta x) - x_{s_i}\right)^2 + \left((y_o + \Delta y) - y_{s_i}\right)^2 + \left(z_o - z_{s_i}\right)^2}}{c} + \frac{\sqrt{\left(x_{r_j} - (x_o + \Delta x)\right)^2 + \left(y_{r_j} - (y_o + \Delta y)\right)^2 + \left(z_{r_j} - z_o\right)^2}}{c} \quad \text{(A.2.2)}$$

$$
\begin{aligned}
t_{s_i r_j}^{(a_3)}\left(\mathbf{x}_o\right) = &\frac{\sqrt{\left(\left(x_o+\Delta x\right)-x_{s_i}\right)^2+\left(\left(y_o-\Delta y\right)-y_{s_i}\right)^2+\left(z_o-z_{s_i}\right)^2}}{c} \\
&+\frac{\sqrt{\left(x_{r_j}-\left(x_o+\Delta x\right)\right)^2+\left(y_{r_j}-\left(y_o-\Delta y\right)\right)^2+\left(z_{r_j}-z_o\right)^2}}{c}
\end{aligned}
\tag{A.2.3}
$$

$$
\begin{aligned}
t_{s_i r_j}^{(a_4)}\left(\mathbf{x}_o\right) = &\frac{\sqrt{\left(\left(x_o-\Delta x\right)-x_{s_i}\right)^2+\left(\left(y_o-\Delta y\right)-y_{s_i}\right)^2+\left(z_o-z_{s_i}\right)^2}}{c} \\
&+\frac{\sqrt{\left(x_{r_j}-\left(x_o-\Delta x\right)\right)^2+\left(y_{r_j}-\left(y_o-\Delta y\right)\right)^2+\left(z_{r_j}-z_o\right)^2}}{c}
\end{aligned}
\tag{A.2.4}
$$

A.3 Fisher information matrix entries

In order to compute the CRLB, the entries of the Fisher information matrix (FIM) are required as shown in (15). The partial derivatives of $\tilde{h}_{s_i r_j}\left(t, \mathbf{x}_o\right)$ with respect to $\mathbf{x}_{o,k}$, $k = 1, 2, 3$, are expanded in (A.3.1), (A.3.2), and (A.3.3) respectively where $\mathbf{x}_o = \left[x_o, y_o, z_o\right]^T$.

$$
\begin{aligned}
\frac{\partial \tilde{h}_{s_i r_j}\left(t, \mathbf{x}_o\right)}{\partial x_o} = &-\eta_{s_i r_j}\left(\mathbf{x}_o\right)\frac{\partial}{\partial x_o}\left\{t_{s_i r_j}\left(\mathbf{x}_o\right)\right\} e^{-\frac{\left(t-t_{s_i r_j}\left(\mathbf{x}_o\right)\right)}{\tau_{s_i r_j}\left(\mathbf{x}_o\right)}} \delta_{\sigma_a,\varepsilon}\left(t-t_{s_i r_j}\left(\mathbf{x}_o\right)\right) \\
&+\frac{\partial}{\partial x_o}\left\{\eta_{s_i r_j}\left(\mathbf{x}_o\right) e^{-\frac{\left(t-t_{s_i r_j}\left(\mathbf{x}_o\right)\right)}{\tau_{s_i r_j}\left(\mathbf{x}_o\right)}}\right\} u_{\sigma_a,\varepsilon}\left(t-t_{s_i r_j}\left(\mathbf{x}_o\right)\right)
\end{aligned}
\tag{A.3.1}
$$

$$
\begin{aligned}
\frac{\partial \tilde{h}_{s_i r_j}\left(t, \mathbf{x}_o\right)}{\partial y_o} = &-\eta_{s_i r_j}\left(\mathbf{x}_o\right)\frac{\partial}{\partial y_o}\left\{t_{s_i r_j}\left(\mathbf{x}_o\right)\right\} e^{-\frac{\left(t-t_{s_i r_j}\left(\mathbf{x}_o\right)\right)}{\tau_{s_i r_j}\left(\mathbf{x}_o\right)}} \delta_{\sigma_a,\varepsilon}\left(t-t_{s_i r_j}\left(\mathbf{x}_o\right)\right) \\
&+\frac{\partial}{\partial y_o}\left\{\eta_{s_i r_j}\left(\mathbf{x}_o\right) e^{-\frac{\left(t-t_{s_i r_j}\left(\mathbf{x}_o\right)\right)}{\tau_{s_i r_j}\left(\mathbf{x}_o\right)}}\right\} u_{\sigma_a,\varepsilon}\left(t-t_{s_i r_j}\left(\mathbf{x}_o\right)\right)
\end{aligned}
\tag{A.3.2}
$$

$$\begin{aligned}\frac{\partial \tilde{h}_{s_i r_j}(t,\mathbf{x}_o)}{\partial z_o} &= -\eta_{s_i r_j}(\mathbf{x}_o)\frac{\partial}{\partial z_o}\left\{t_{s_i r_j}(\mathbf{x}_o)\right\} e^{-\frac{\left(t-t_{s_i r_j}(\mathbf{x}_o)\right)}{\tau_{s_i r_j}(\mathbf{x}_o)}}\delta_{\sigma_a,\varepsilon}\left(t-t_{s_i r_j}(\mathbf{x}_o)\right)\\ &\quad+\frac{\partial}{\partial z_o}\left\{\eta_{s_i r_j}(\mathbf{x}_o)\,e^{-\frac{\left(t-t_{s_i r_j}(\mathbf{x}_o)\right)}{\tau_{s_i r_j}(\mathbf{x}_o)}}\right\}u_{\sigma_a,\varepsilon}\left(t-t_{s_i r_j}(\mathbf{x}_o)\right)\end{aligned}\tag{A.3.3}$$

A.3.1 Partial derivatives with respect to x_o

The partial derivative of first term of (A.3.1) with respect to x_o is given in (A.3.4) and the second term can be expanded using the product rule of derivatives as shown in (A.3.5). The partial derivative $\frac{\partial \tau_{s_i r_j}(\mathbf{x}_o)}{\partial x_o}$ can be expanded as shown in (A.3.6). The partial derivative with respect to x_o of time delay from the corners a_1, a_2, a_3, and a_4 are derived in (A.3.7), (A.3.8), (A.3.9), and (A.3.10) respectively.

$$\frac{\partial}{\partial x_o}\left\{t_{s_i r_j}(\mathbf{x}_o)\right\} = \left(\frac{x_o - x_{s_i}}{c\left\|\mathbf{x}_o-\mathbf{x}_i^{(s)}\right\|_2} - \frac{x_{r_j}-x_o}{c\left\|\mathbf{x}_j^{(r)}-\mathbf{x}_o\right\|_2}\right)\tag{A.3.4}$$

$$\begin{aligned}&\frac{\partial}{\partial x_o}\left\{\eta_{s_i r_j}(\mathbf{x}_o)\,e^{-\frac{\left(t-t_{s_i r_j}(\mathbf{x}_o)\right)}{\tau_{s_i r_j}(\mathbf{x}_o)}}\right\}\\ &\quad=\eta_{s_i r_j}(\mathbf{x}_o)\,e^{-\frac{\left(t-t_{s_i r_j}(\mathbf{x}_o)\right)}{\tau_{s_i r_j}(\mathbf{x}_o)}}\left(\frac{\tau_{s_i r_j}(\mathbf{x}_o)\frac{\partial t_{s_i r_j}(\mathbf{x}_o)}{\partial x_o}-\left(t_{s_i r_j}(\mathbf{x}_o)-t\right)\frac{\partial \tau_{s_i r_j}(\mathbf{x}_o)}{\partial x_o}}{\left(\tau_{s_i r_j}(\mathbf{x}_o)\right)^2}\right)\\ &\qquad+\eta_{s_i r_j}(\mathbf{x}_o)\,e^{-\frac{\left(t-t_{s_i r_j}(\mathbf{x}_o)\right)}{\tau_{s_i r_j}(\mathbf{x}_o)}}\left(\frac{4\left(x_{r_j}-x_o\right)}{\left\|\mathbf{x}_j^{(r)}-\mathbf{x}_o\right\|_2^2}-\frac{(m+3)\left(x_o-x_{s_i}\right)}{\left\|\mathbf{x}_o-\mathbf{x}_i^{(s)}\right\|_2^2}\right)\end{aligned}\tag{A.3.5}$$

$$\begin{aligned}\frac{\partial \tau_{s_ir_j}(\mathbf{x}_o)}{\partial x_o} &= \frac{1}{4\tau_{s_ir_j}(\mathbf{x}_o)}\\ &\times\Bigg\{\left(t^{(a_1)}_{s_ir_j}(\mathbf{x}_o)-t_{s_ir_j}(\mathbf{x}_o)\right)\left(\frac{\partial t^{(a_1)}_{s_ir_j}(\mathbf{x}_o)}{\partial x_o}-\frac{\partial t_{s_ir_j}(\mathbf{x}_o)}{\partial x_o}\right)\\ &+\left(t^{(a_2)}_{s_ir_j}(\mathbf{x}_o)-t_{s_ir_j}(\mathbf{x}_o)\right)\left(\frac{\partial t^{(a_2)}_{s_ir_j}(\mathbf{x}_o)}{\partial x_o}-\frac{\partial t_{s_ir_j}(\mathbf{x}_o)}{\partial x_o}\right)\\ &+\left(t^{(a_3)}_{s_ir_j}(\mathbf{x}_o)-t_{s_ir_j}(\mathbf{x}_o)\right)\left(\frac{\partial t^{(a_3)}_{s_ir_j}(\mathbf{x}_o)}{\partial x_o}-\frac{\partial t_{s_ir_j}(\mathbf{x}_o)}{\partial x_o}\right)\\ &+\left(t^{(a_4)}_{s_ir_j}(\mathbf{x}_o)-t_{s_ir_j}(\mathbf{x}_o)\right)\left(\frac{\partial t^{(a_4)}_{s_ir_j}(\mathbf{x}_o)}{\partial x_o}-\frac{\partial t_{s_ir_j}(\mathbf{x}_o)}{\partial x_o}\right)\Bigg\}\end{aligned} \tag{A.3.6}$$

$$\begin{aligned}\frac{\partial t^{(a_1)}_{s_ir_j}(\mathbf{x}_o)}{\partial x_o} &= \frac{c^{-1}\left((x_o-\Delta x)-x_{s_i}\right)}{\sqrt{\left((x_o-\Delta x)-x_{s_i}\right)^2+\left((y_o+\Delta y)-y_{s_i}\right)^2+\left(z_o-z_{s_i}\right)^2}}\\ &-\frac{c^{-1}\left(x_{r_j}-(x_o-\Delta x)\right)}{\sqrt{\left(x_{r_j}-(x_o-\Delta x)\right)^2+\left(y_{r_j}-(y_o+\Delta y)\right)^2+\left(z_{r_j}-z_o\right)^2}}\end{aligned} \tag{A.3.7}$$

$$\begin{aligned}\frac{\partial t^{(a_2)}_{s_ir_j}(\mathbf{x}_o)}{\partial x_o} &= \frac{c^{-1}\left((x_o+\Delta x)-x_{s_i}\right)}{\sqrt{\left((x_o+\Delta x)-x_{s_i}\right)^2+\left((y_o+\Delta y)-y_{s_i}\right)^2+\left(z_o-z_{s_i}\right)^2}}\\ &-\frac{c^{-1}\left(x_{r_j}-(x_o+\Delta x)\right)}{\sqrt{\left(x_{r_j}-(x_o+\Delta x)\right)^2+\left(y_{r_j}-(y_o+\Delta y)\right)^2+\left(z_{r_j}-z_o\right)^2}}\end{aligned} \tag{A.3.8}$$

$$\begin{aligned}\frac{\partial t^{(a_3)}_{s_ir_j}(\mathbf{x}_o)}{\partial x_o} &= \frac{c^{-1}\left((x_o+\Delta x)-x_{s_i}\right)}{\sqrt{\left((x_o+\Delta x)-x_{s_i}\right)^2+\left((y_o-\Delta y)-y_{s_i}\right)^2+\left(z_o-z_{s_i}\right)^2}}\\ &-\frac{c^{-1}\left(x_{r_j}-(x_o+\Delta x)\right)}{\sqrt{\left(x_{r_j}-(x_o+\Delta x)\right)^2+\left(y_{r_j}-(y_o-\Delta y)\right)^2+\left(z_{r_j}-z_o\right)^2}}\end{aligned} \tag{A.3.9}$$

$$\begin{aligned}\frac{\partial t^{(a_4)}_{s_i r_j}\left(\mathbf{x}_o\right)}{\partial x_o} &= \frac{c^{-1}\left(\left(x_o-\Delta x\right)-x_{s_i}\right)}{\sqrt{\left(\left(x_o-\Delta x\right)-x_{s_i}\right)^2+\left(\left(y_o-\Delta y\right)-y_{s_i}\right)^2+\left(z_o-z_{s_i}\right)^2}} \\ &\quad - \frac{c^{-1}\left(x_{r_j}-\left(x_o-\Delta x\right)\right)}{\sqrt{\left(x_{r_j}-\left(x_o-\Delta x\right)\right)^2+\left(y_{r_j}-\left(y_o-\Delta y\right)\right)^2+\left(z_{r_j}-z_o\right)^2}} \qquad \text{(A.3.10)}\end{aligned}$$

A.3.2 Partial derivatives with respect to y_o

Similarly, (A.3.11) gives the partial derivative of the first term of (A.3.2) with respect to y_o and (A.3.12) gives the expansion of the second term. The partial derivative $\frac{\partial \tau_{s_i r_j,m}(\mathbf{x}_o)}{\partial y_o}$ is given in (A.3.13) while the partial derivatives with respect to y_o of the time delay from the corners is given in in (A.3.14), (A.3.15), (A.3.16), and (A.3.17).

$$\frac{\partial}{\partial y_o}\left\{t_{s_i r_j}\left(\mathbf{x}_o\right)\right\} = \left(\frac{y_o-y_{s_i}}{c\left\|\mathbf{x}_o-\mathbf{x}_i^{(s)}\right\|_2} - \frac{y_{r_j}-y_o}{c\left\|\mathbf{x}_j^{(r)}-\mathbf{x}_o\right\|_2}\right) \qquad \text{(A.3.11)}$$

$$\begin{aligned}&\frac{\partial}{\partial y_o}\left\{\eta_{s_i r_j}\left(\mathbf{x}_o\right)e^{-\frac{\left(t-t_{s_i r_j}(\mathbf{x}_o)\right)}{\tau_{s_i r_j}(\mathbf{x}_o)}}\right\} \\ &\quad = \eta_{s_i r_j}\left(\mathbf{x}_o\right)e^{-\frac{\left(t-t_{s_i r_j}(\mathbf{x}_o)\right)}{\tau_{s_i r_j}(\mathbf{x}_o)}}\left(\frac{\tau_{s_i r_j}\left(\mathbf{x}_o\right)\frac{\partial t_{s_i r_j}(\mathbf{x}_o)}{\partial y_o}-\left(t_{s_i r_j}\left(\mathbf{x}_o\right)-t\right)\frac{\partial \tau_{s_i r_j}(\mathbf{x}_o)}{\partial y_o}}{\left(\tau_{s_i r_j}\left(\mathbf{x}_o\right)\right)^2}\right) \\ &\qquad + \eta_{s_i r_j}\left(\mathbf{x}_o\right)e^{-\frac{\left(t-t_{s_i r_j}(\mathbf{x}_o)\right)}{\tau_{s_i r_j}(\mathbf{x}_o)}}\left(\frac{4\left(y_{r_j}-y_o\right)}{\left\|\mathbf{x}_j^{(r)}-\mathbf{x}_o\right\|_2^2}-\frac{(m+3)\left(y_o-y_{s_i}\right)}{\left\|\mathbf{x}_o-\mathbf{x}_i^{(s)}\right\|_2^2}\right) \qquad \text{(A.3.12)}\end{aligned}$$

$$
\begin{aligned}
\frac{\partial \tau_{s_i r_j}(\mathbf{x}_o)}{\partial y_o} &= \frac{1}{4\tau_{s_i r_j}(\mathbf{x}_o)} \\
&\quad \times \Bigg\{ \left(t^{(a_1)}_{s_i r_j}(\mathbf{x}_o) - t_{s_i r_j}(\mathbf{x}_o)\right)\left(\frac{\partial t^{(a_1)}_{s_i r_j}(\mathbf{x}_o)}{\partial y_o} - \frac{\partial t_{s_i r_j}(\mathbf{x}_o)}{\partial y_o}\right) \\
&\quad + \left(t^{(a_2)}_{s_i r_j}(\mathbf{x}_o) - t_{s_i r_j}(\mathbf{x}_o)\right)\left(\frac{\partial t^{(a_2)}_{s_i r_j}(\mathbf{x}_o)}{\partial y_o} - \frac{\partial t_{s_i r_j}(\mathbf{x}_o)}{\partial y_o}\right) \\
&\quad + \left(t^{(a_3)}_{s_i r_j}(\mathbf{x}_o) - t_{s_i r_j}(\mathbf{x}_o)\right)\left(\frac{\partial t^{(a_3)}_{s_i r_j}(\mathbf{x}_o)}{\partial y_o} - \frac{\partial t_{s_i r_j}(\mathbf{x}_o)}{\partial y_o}\right) \\
&\quad + \left(t^{(a_4)}_{s_i r_j}(\mathbf{x}_o) - t_{s_i r_j}(\mathbf{x}_o)\right)\left(\frac{\partial t^{(a_4)}_{s_i r_j}(\mathbf{x}_o)}{\partial y_o} - \frac{\partial t_{s_i r_j}(\mathbf{x}_o)}{\partial y_o}\right)\Bigg\}
\end{aligned} \tag{A.3.13}
$$

$$
\begin{aligned}
\frac{\partial t^{(a_1)}_{s_i r_j}(\mathbf{x}_o)}{\partial y_o} &= \frac{c^{-1}\left((y_o + \Delta y) - y_{s_i}\right)}{\sqrt{\left((x_o - \Delta x) - x_{s_i}\right)^2 + \left((y_o + \Delta y) - y_{s_i}\right)^2 + \left(z_o - z_{s_i}\right)^2}} \\
&\quad - \frac{c^{-1}\left(y_{r_j} - (y_o + \Delta y)\right)}{\sqrt{\left(x_{r_j} - (x_o - \Delta x)\right)^2 + \left(y_{r_j} - (y_o + \Delta y)\right)^2 + \left(z_{r_j} - z_o\right)^2}}
\end{aligned} \tag{A.3.14}
$$

$$
\begin{aligned}
\frac{\partial t^{(a_2)}_{s_i r_j}(\mathbf{x}_o)}{\partial y_o} &= \frac{c^{-1}\left((y_o + \Delta y) - y_{s_i}\right)}{\sqrt{\left((x_o + \Delta x) - x_{s_i}\right)^2 + \left((y_o + \Delta y) - y_{s_i}\right)^2 + \left(z_o - z_{s_i}\right)^2}} \\
&\quad - \frac{c^{-1}\left(y_{r_j} - (y_o + \Delta y)\right)}{\sqrt{\left(x_{r_j} - (x_o + \Delta x)\right)^2 + \left(y_{r_j} - (y_o + \Delta y)\right)^2 + \left(z_{r_j} - z_o\right)^2}}
\end{aligned} \tag{A.3.15}
$$

$$
\begin{aligned}
\frac{\partial t^{(a_3)}_{s_i r_j}(\mathbf{x}_o)}{\partial y_o} &= \frac{c^{-1}\left((y_o - \Delta y) - y_{s_i}\right)}{\sqrt{\left((x_o + \Delta x) - x_{s_i}\right)^2 + \left((y_o - \Delta y) - y_{s_i}\right)^2 + \left(z_o - z_{s_i}\right)^2}} \\
&\quad - \frac{c^{-1}\left(y_{r_j} - (y_o - \Delta y)\right)}{\sqrt{\left(x_{r_j} - (x_o + \Delta x)\right)^2 + \left(y_{r_j} - (y_o - \Delta y)\right)^2 + \left(z_{r_j} - z_o\right)^2}}
\end{aligned} \tag{A.3.16}
$$

$$\frac{\partial t_{s_i r_j}^{(a_4)}(\mathbf{x}_o)}{\partial y_o} = \frac{c^{-1}\left((y_o-\Delta y)-y_{s_i}\right)}{\sqrt{\left((x_o-\Delta x)-x_{s_i}\right)^2+\left((y_o-\Delta y)-y_{s_i}\right)^2+\left(z_o-z_{s_i}\right)^2}} - \frac{c^{-1}\left(y_{r_j}-(y_o-\Delta y)\right)}{\sqrt{\left(x_{r_j}-(x_o-\Delta x)\right)^2+\left(y_{r_j}-(y_o-\Delta y)\right)^2+\left(z_{r_j}-z_o\right)^2}} \tag{A.3.17}$$

A.3.3 Partial derivatives with respect to z_o

Equation (A.3.18) gives the partial derivative of the first term of (A.3.3) with respect to z_o. The second term of (A.3.3) is expanded in (A.3.19). The partial derivative $\frac{\partial \tau_{s_i r_j,m}(\mathbf{x}_o)}{\partial z_o}$ is expanded in (A.3.20) while the partial derivatives with respect to z_o for times delays from the corners of $\mathcal{S}_o$ are given in (A.3.21), (A.3.22), (A.3.23), and (A.3.24) respectively.

$$\frac{\partial}{\partial z_o}\left\{t_{s_i r_j}(\mathbf{x}_o)\right\} = \left(\frac{z_o-z_{s_i}}{c\left\|\mathbf{x}_o-\mathbf{x}_i^{(s)}\right\|_2} - \frac{z_{r_j}-z_o}{c\left\|\mathbf{x}_j^{(r)}-\mathbf{x}_o\right\|_2}\right) \tag{A.3.18}$$

$$\begin{aligned}\frac{\partial}{\partial z_o}&\left\{\eta_{s_i r_j}(\mathbf{x}_o)\, e^{-\frac{\left(t-t_{s_i r_j}(\mathbf{x}_o)\right)}{\tau_{s_i r_j}(\mathbf{x}_o)}}\right\}\\ &= \eta_{s_i r_j}(\mathbf{x}_o)\, e^{-\frac{\left(t-t_{s_i r_j}(\mathbf{x}_o)\right)}{\tau_{s_i r_j}(\mathbf{x}_o)}} \left(\frac{\tau_{s_i r_j}(\mathbf{x}_o)\frac{\partial t_{s_i r_j}(\mathbf{x}_o)}{\partial z_o} - \left(t_{s_i r_j}(\mathbf{x}_o)-t\right)\frac{\partial \tau_{s_i r_j}(\mathbf{x}_o)}{\partial z_o}}{\left(\tau_{s_i r_j}(\mathbf{x}_o)\right)^2}\right)\\ +\eta_{s_i r_j}(\mathbf{x}_o)\, e^{-\frac{\left(t-t_{s_i r_j}(\mathbf{x}_o)\right)}{\tau_{s_i r_j}(\mathbf{x}_o)}}&\left(-\frac{2}{z_{r_j}-z_o} - \frac{(m+1)}{z_{s_i}-z_o} + \frac{4\left(z_{r_j}-z_o\right)}{\left\|\mathbf{x}_j^{(r)}-\mathbf{x}_o\right\|_2^2} - \frac{(m+3)\left(z_o-z_{s_i}\right)}{\left\|\mathbf{x}_o-\mathbf{x}_i^{(s)}\right\|_2^2}\right)\end{aligned} \tag{A.3.19}$$

$$\begin{aligned}\frac{\partial \tau_{s_i r_j}(\mathbf{x}_o)}{\partial z_o} &= \frac{1}{4\tau_{s_i r_j}(\mathbf{x}_o)} \\ &\times \Bigg\{ \left(t^{(a_1)}_{s_i r_j}(\mathbf{x}_o) - t_{s_i r_j}(\mathbf{x}_o)\right)\left(\frac{\partial t^{(a_1)}_{s_i r_j}(\mathbf{x}_o)}{\partial z_o} - \frac{\partial t_{s_i r_j}(\mathbf{x}_o)}{\partial z_o}\right) \\ &+ \left(t^{(a_2)}_{s_i r_j}(\mathbf{x}_o) - t_{s_i r_j}(\mathbf{x}_o)\right)\left(\frac{\partial t^{(a_2)}_{s_i r_j}(\mathbf{x}_o)}{\partial z_o} - \frac{\partial t_{s_i r_j}(\mathbf{x}_o)}{\partial z_o}\right) \\ &+ \left(t^{(a_3)}_{s_i r_j}(\mathbf{x}_o) - t_{s_i r_j}(\mathbf{x}_o)\right)\left(\frac{\partial t^{(a_3)}_{s_i r_j}(\mathbf{x}_o)}{\partial z_o} - \frac{\partial t_{s_i r_j}(\mathbf{x}_o)}{\partial z_o}\right) \\ &+ \left(t^{(a_4)}_{s_i r_j}(\mathbf{x}_o) - t_{s_i r_j}(\mathbf{x}_o)\right)\left(\frac{\partial t^{(a_4)}_{s_i r_j}(\mathbf{x}_o)}{\partial z_o} - \frac{\partial t_{s_i r_j}(\mathbf{x}_o)}{\partial z_o}\right)\Bigg\} \end{aligned} \tag{A.3.20}$$

$$\begin{aligned}\frac{\partial t^{(a_1)}_{s_i r_j}(\mathbf{x}_o)}{\partial z_o} &= \frac{c^{-1}(z_o - z_{s_i})}{\sqrt{((x_o - \Delta x) - x_{s_i})^2 + ((y_o + \Delta y) - y_{s_i})^2 + (z_o - z_{s_i})^2}} \\ &- \frac{c^{-1}\left(z_{r_j} - z_o\right)}{\sqrt{\left(x_{r_j} - (x_o - \Delta x)\right)^2 + \left(y_{r_j} - (y_o + \Delta y)\right)^2 + \left(z_{r_j} - z_o\right)^2}} \end{aligned} \tag{A.3.21}$$

$$\begin{aligned}\frac{\partial t^{(a_2)}_{s_i r_j}(\mathbf{x}_o)}{\partial z_o} &= \frac{c^{-1}(z_o - z_{s_i})}{\sqrt{((x_o + \Delta x) - x_{s_i})^2 + ((y_o + \Delta y) - y_{s_i})^2 + (z_o - z_{s_i})^2}} \\ &- \frac{c^{-1}\left(z_{r_j} - z_o\right)}{\sqrt{\left(x_{r_j} - (x_o + \Delta x)\right)^2 + \left(y_{r_j} - (y_o + \Delta y)\right)^2 + \left(z_{r_j} - z_o\right)^2}} \end{aligned} \tag{A.3.22}$$

$$\begin{aligned}\frac{\partial t^{(a_3)}_{s_i r_j}(\mathbf{x}_o)}{\partial z_o} &= \frac{c^{-1}(z_o - z_{s_i})}{\sqrt{((x_o + \Delta x) - x_{s_i})^2 + ((y_o - \Delta y) - y_{s_i})^2 + (z_o - z_{s_i})^2}} \\ &- \frac{c^{-1}\left(z_{r_j} - z_o\right)}{\sqrt{\left(x_{r_j} - (x_o + \Delta x)\right)^2 + \left(y_{r_j} - (y_o - \Delta y)\right)^2 + \left(z_{r_j} - z_o\right)^2}} \end{aligned} \tag{A.3.23}$$

$$\begin{aligned}\frac{\partial t_{s_i r_j}^{(a_4)}\left(\mathbf{x}_o\right)}{\partial z_o} &= \frac{c^{-1}\left(z_o - z_{s_i}\right)}{\sqrt{\left(\left(x_o - \Delta x\right) - x_{s_i}\right)^2 + \left(\left(y_o - \Delta y\right) - y_{s_i}\right)^2 + \left(z_o - z_{s_i}\right)^2}} \\ &\quad - \frac{c^{-1}\left(z_{r_j} - z_o\right)}{\sqrt{\left(x_{r_j} - \left(x_o - \Delta x\right)\right)^2 + \left(y_{r_j} - \left(y_o - \Delta y\right)\right)^2 + \left(z_{r_j} - z_o\right)^2}} \qquad \text{(A.3.24)}\end{aligned}$$

Bibliography

[1] N. Samama, *Global Positioning: Technologies and Performance.* Wiley-Interscience, 605 Third Avenue New York, NY, United States, 2008.

[2] A. El-Rabbany, "Introduction to GPS: The global positioning system." Artech, 2006.

[3] A. K. Dubey, "Global positioning system-the newest utility," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 15, no. 10, pp. 89–95, oct 2000.

[4] F. van Diggelen and P. Enge, "The world's first GPS MOOC and worldwide laboratory using smartphones," in *Proc. 28th Int. Tech. Meet. Satell. Div. Inst. Navig. (ION GNSS+ 2015)*, Tampa, Florida, 2015, pp. 361 – 369.

[5] A. Raza, S. Hameed, and T. Macintyre, "Global positioning system - Working and its applications," in *Innov. Adv. Tech. Syst. Comput. Sci. Softw. Eng.*, 2008, pp. 448–453.

[6] R. Watson, G. Lachapelle, R. Klukas, S. Turunen, S. Pietilä, and I. Halivaara, "Investigating GPS signals indoors with extreme high-sensitivity detection techniques," *J. Inst. Navig.*, vol. 52, no. 4, pp. 199–213, 2005.

[7] S. He and S. H. G. Chan, "Wi-Fi fingerprint-based indoor positioning: Recent advances and comparisons," *IEEE Commun. Surv. Tutorials*, vol. 18, no. 1, pp. 466–490, 2016.

[8] P. Davidson and R. Piche, "A survey of selected indoor positioning methods for smartphones," *IEEE Commun. Surv. Tutorials*, vol. 19, no. 2, pp. 1347–1370, 2016.

[9] A. Yassin, Y. Nasser, M. Awad, A. Al-Dubai, R. Liu, C. Yuen, R. Raulefs, and E. Aboutanios, "Recent advances in indoor localization: A survey on theoretical approaches and applications," *IEEE Commun. Surv. Tutorials*, vol. 19, no. 2, pp. 1327–1346, 2017.

[10] F. Zafari, A. Gkelias, and K. K. Leung, "A survey of indoor localization systems and technologies," *IEEE Commun. Surv. Tutorials*, vol. 21, no. 3, pp. 2568–2599, 2019.

[11] S. Gezici, Z. Tian, G. B. Giannakis, H. Kobayashi, A. F. Molisch, H. V. Poor, and Z. Sahinoglu, "Localization via ultra-wideband radios: A look at positioning aspects of future sensor networks," *IEEE Signal Process. Mag.*, vol. 22, no. 4, pp. 70–84, 2005.

[12] Z. Song, G. Jiang, and C. Huang, "A survey on indoor positioning technologies," in *Commun. Comput. Inf. Sci.* Springer Berlin Heidelberg, 2011, pp. 198–206.

[13] IEEE, "Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," *IEEE Std 802.11-2012 (Revision IEEE Std*

802.11-2007), vol. 2012, no. March, p. 2793, 2012. [Online]. Available: http://ieeexplore.ieee.org/servlet/opac?punumber=6178209

[14] S. Sorour, Y. Lostanlen, S. Valaee, and K. Majeed, "Joint indoor localization and radio map construction with limited deployment load," *IEEE Trans. Mob. Comput.*, vol. 14, no. 5, pp. 1031–1043, 2015.

[15] K. Majeed, S. Sorour, T. Y. Al-Naffouri, and S. Valaee, "Indoor localization using unsupervised manifold alignment with geometry perturbation," in *2014 IEEE Wirel. Commun. Netw. Conf.*, Istanbul, 2014, pp. 2952–2957.

[16] ——, "Indoor localization and radio map estimation using unsupervised manifold alignment with geometry perturbation," *IEEE Trans. Mob. Comput.*, vol. 15, no. 11, pp. 2794–2808, 2016.

[17] S. Li, M. Hedley, K. Bengston, D. Humphrey, M. Johnson, and W. Ni, "Passive localization of standard WiFi devices," *IEEE Syst. J.*, vol. 13, no. 4, pp. 3929–3932, 2019.

[18] ——, "TDOA-based passive localization of standard WiFi devices," in *Ubiquitous Positioning, Indoor Navig. Locat. Serv.*, 2018, pp. 1–5.

[19] H. X. Chen, B. J. Hu, L. L. Zheng, and Z. H. Wei, "An accurate AoA estimation approach for indoor localization using commodity Wi-Fi devices," in *2018 IEEE Int. Conf. Signal Process. Commun. Comput. ICSPCC 2018*, 2018, pp. 1–5.

[20] J. Krumm and J. Platt, “Minimizing calibration effort for an indoor 802.11 device location measurement system,” Tech. Rep., 2003. [Online]. Available: https://www.microsoft.com/en-us/research/publication/minimizing-calibration-effort-for-an-indoor-802-11-device-location-measurement-system/

[21] Apple Inc., “iBeacon.” [Online]. Available: https://developer.apple.com/ibeacon/

[22] T. Jingwangsa, S. Soonjun, and P. Chemtanomwong, “Comparison between innovative approaches of RFID based localization using fingerprinting techniques for outdoor and indoor environments,” in *Int. Conf. Adv. Commun. Technol. ICACT*, vol. 2, 2010, pp. 1511–1515.

[23] I. Besic, Z. Avdagic, and K. Hodzic, “RFID based indoors test setup for visually impaired,” in *ELECO 2019 - 11th Int. Conf. Electr. Electron. Eng.*, no. 1, 2019, pp. 465–469.

[24] Z. Yang, C. Wu, Z. Zhou, X. Zhang, X. Wang, and Y. Liu, “Mobility increases localizability: A survey on wireless indoor localization using inertial sensors,” *ACM Comput. Surv.*, vol. 47, no. 3, 2015.

[25] F. W. Gobana, “Survey of inertial/magnetic sensors based pedestrian dead reckoning by multi-sensor fusion method,” in *9th Int. Conf. Inf. Commun. Technol. Converg. ICT Converg. Powered by Smart Intell. ICTC 2018*, 2018, pp. 1327–1334.

[26] W. Huang, Y. Xiong, X. Y. Li, H. Lin, X. Mao, P. Yang, Y. Liu, and X. Wang, “Swadloon: Direction finding and indoor localization using acoustic signal by

shaking smartphones," *IEEE Trans. Mob. Comput.*, vol. 14, no. 10, pp. 2145–2157, 2015.

[27] M. Hazas and A. Hopper, "Broadband ultrasonic location systems for improved indoor positioning," *IEEE Trans. Mob. Comput.*, vol. 5, no. 5, pp. 536–547, 2006.

[28] F. Ijaz, H. K. Yang, A. W. Ahmad, and C. Lee, "Indoor positioning: A review of indoor ultrasonic positioning systems," in *Int. Conf. Adv. Commun. Technol. ICACT*. IEEE, 2013, pp. 1146–1150.

[29] P. Palojarvi, M. Kari, and J. Kostamovaara, "Integrated time-of-flight laser radar," *IEEE Trans. Instrum. Meas.*, vol. 46, no. 4, pp. 996–999, 1997.

[30] M. J. Gallant and J. A. Marshall, "Two-dimensional axis mapping using LiDAR," *IEEE Trans. Robot.*, vol. 32, no. 1, pp. 150–160, 2016.

[31] Q. Wu, K. Sun, W. Zhang, C. Huang, and X. Wu, "Visual and LiDAR-based for the mobile 3D mapping," in *2016 IEEE Int. Conf. Robot. Biomimetics, ROBIO 2016*, no. 2, 2016, pp. 1522–1527.

[32] H. Griffiths, "Early history of bistatic radar," in *2016 Eur. Radar Conf. EuRAD 2016*. EuMA, 2016, pp. 253–257.

[33] G. Ajay Kumar, A. K. Patil, R. Patil, S. S. Park, and Y. H. Chai, "A LiDAR and IMU integrated indoor navigation system for UAVs and its application in real-time pipeline classification," *Sensors (Switzerland)*, vol. 17, no. 6, 2017.

[34] T. H. Kim and T. H. Park, “Placement optimization of multiple lidar sensors for autonomous vehicles,” *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 5, pp. 2139–2145, 2020.

[35] Y. Chen, J. Tang, C. Jiang, L. Zhu, M. Lehtomäki, H. Kaartinen, R. Kaijaluoto, Y. Wang, J. Hyyppä, H. Hyyppä, H. Zhou, L. Pei, and R. Chen, “The accuracy comparison of three simultaneous localization and mapping (SLAM)-based indoor mapping technologies,” *Sensors (Switzerland)*, vol. 18, no. 10, 2018.

[36] Z. Xuexi, L. Guokun, F. Genping, X. Dongliang, and L. Shiliu, “SLAM algorithm analysis of mobile robot based on LiDAR,” in *Chinese Control Conf. CCC.* Technical Committee on Control Theory, Chinese Association of Automation, 2019, pp. 4739–4745.

[37] H. S. Liu and G. Pang, “Positioning beacon system using digital camera and LEDs,” *IEEE Trans. Veh. Technol.*, vol. 52, no. 2, pp. 406–419, 2003.

[38] C. Zhang and X. Zhang, “Demo: LiTell: Indoor localization using unmodified light fixtures,” *Proc. Annu. Int. Conf. Mob. Comput. Networking, MOBICOM*, vol. 0, no. 1, pp. 481–482, 2016.

[39] G. K. H. Pang and H. H. S. Liu, “LED location beacon system based on processing of digital images,” *IEEE Trans. Intell. Transp. Syst.*, vol. 2, no. 3, pp. 135–150, 2001.

[40] P. Chavez-Burbano, V. Guerra, J. Rabadan, C. Jurado-Verdu, and R. Perez-Jimenez, “Novel indoor localization system using optical camera communication,” in *2018 11th Int. Symp. Commun. Syst. Networks Digit. Signal Process.*

CSNDSP 2018. IEEE, 2018, pp. 1–5.

[41] C. Zhang and X. Zhang, “Visible light localization using conventional light fixtures and smartphones,” *IEEE Trans. Mob. Comput.*, vol. 18, no. 12, pp. 2968–2983, 2019.

[42] Interact Lighting, “Indoor Navigation,” 2021. [Online]. Available: https://www.interact-lighting.com/global/what-is-possible/interact-office/indoor-navigation

[43] P. Speciale, J. L. Schonberger, S. B. Kang, S. N. Sinha, and M. Pollefeys, “Privacy preserving image-based localization,” in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 5488–5498.

[44] C. Debeunne and D. Vivet, “A review of visual-LiDAR fusion based simultaneous localization and mapping,” *Sensors (Switzerland)*, vol. 20, no. 7, 2020.

[45] J. M. Kahn and J. R. Barry, “Wireless infrared communications,” *Proc. IEEE*, vol. 9219, no. 97, pp. 265–298, 1997.

[46] T. Do and M. Yoo, “An in-depth survey of visible light communication based positioning systems,” *Sensors*, vol. 16, no. 5, p. 678, 2016.

[47] Y. Zhuang, L. Hua, L. Qi, J. Yang, P. Cao, Y. Cao, Y. Wu, J. Thompson, and H. Haas, “A survey of positioning systems using visible LED lights,” *IEEE Commun. Surv. Tutorials*, vol. 20, no. 3, pp. 1963–1988, 2018.

[48] P. H. Pathak, X. Feng, P. Hu, and P. Mohapatra, “Visible light communication, networking, and sensing: A survey, potential and challenges,” *IEEE Commun. Surv. Tutorials*, vol. 17, no. 4, pp. 2047–2077, 2015.

[49] "IEEE standard for local and metropolitan area networks–Part 15.7: Short-range optical wireless communications," *IEEE Std 802.15.7-2018 (Revision IEEE Std 802.15.7-2011)*, pp. 1–407, apr 2019.

[50] "IEEE 802.11™ launches standards amendment project for light communications (LiFi)," 2018. [Online]. Available: https://beyondstandards.ieee.org/ieee-802-11-launches-standards-amendment-project-for-light-communications-lifi/

[51] J. Armstrong, Y. Sekercioglu Ahmet, and A. Neild, "Visible light positioning: A roadmap for international standardization," *IEEE Commun. Mag.*, vol. 51, no. 12, pp. 68–73, 2013.

[52] M. Maheepala, A. Z. Kouzani, and M. A. Joordens, "Light-based indoor positioning systems: A review," *IEEE Sens. J.*, vol. 20, no. 8, pp. 3971–3995, 2020.

[53] J. Singh and U. Raza, "Passive visible light positioning systems: An overview," in *Proc. Work. Light Up IoT*. London: Association for Computing Machinery, 2020, pp. 48–53.

[54] D. Karunatilaka, F. Zafar, V. Kalavally, and R. Parthiban, "LED based indoor visible light communications: State of the art," *IEEE Commun. Surv. Tutorials*, vol. 17, no. 3, pp. 1649–1678, 2015.

[55] Vector Clipart, "Standing person silhouette," 2020. [Online]. Available: https://clipart.me/business-finance/business-people-vector-graphics-36667

[56] Cisco Systems, “Cisco Digital Building Solutions,” 2018. [Online]. Available: https://www.cisco.com/c/en/us/solutions/workforce-experience/digital-building/index.html

[57] G. B. Prince and T. D. Little, “Two-phase framework for indoor positioning systems using visible light,” *Sensors (Switzerland)*, vol. 18, no. 6, 2018.

[58] S. Ayub, B. Honary, S. Kariyawasam, and M. Honary, “Visible light ID system for indoor localization,” in *5th IET Int. Conf. Wireless, Mob. Multimed. Networks (ICWMMN 2013)*, 2013, pp. 5.02–5.02.

[59] A. Danideh and R. A. Sadeghzadeh, “CPW-fed slot antenna for MIMO system applications,” *Indian J. Sci. Technol.*, vol. 6, no. 1, pp. 3872–3875, 2013.

[60] L. Li, P. Hu, C. Peng, G. Shen, and F. Zhao, “Epsilon: A visible light based positioning system,” in *USENIX Symp. Netw. Syst. Des. Implement.*, no. 1, 2014, pp. 331–343.

[61] Z. Zheng, L. Liu, and W. Hu, “Accuracy of ranging based on DMT visible light communication for indoor positioning,” *IEEE Photonics Technol. Lett.*, vol. 29, no. 8, pp. 679–682, 2017.

[62] C. Kottke, S. Nerreter, K.-d. Langer, J. W. Walewski, and J. Vučić, “513 Mbit/s visible light communications link based on DMT-modulation of a white LED,” *J. Light. Technol.*, vol. 28, no. 24, pp. 3512–3518, 2010.

[63] K. Y. Yi, D. Y. Kim, and K. M. Yi, “Development of a localization system based on VLC technique for an indoor environment,” *J. Electr. Eng. Technol.*, vol. 10, no. 1, pp. 436–442, 2015.

[64] X. Zhang, J. Duan, Y. Fu, and A. Shi, “Theoretical accuracy analysis of indoor visible light communication positioning system based on received signal strength indicator,” *J. Light. Technol.*, vol. 32, no. 21, pp. 3578–3584, 2014.

[65] H. Hosseinianfar, M. Noshad, and M. Brandt-Pearce, “Positioning for visible light communication system exploiting multipath reflections,” in *IEEE Int. Conf. Commun.*, 2017.

[66] E. Edwards and S. Hranilovic, “Indoor localization using low-complexity luminaires and ambient light sensors,” in *2016 IEEE Photonics Soc. Summer Top. Meet. Ser.* IEEE, 2016, pp. 154–155.

[67] H. Steendam, T. Q. Wang, and J. Armstrong, “Theoretical lower bound for indoor visible light positioning using received signal strength measurements and an aperture-based receiver,” *J. Light. Technol.*, vol. 35, no. 2, pp. 309–319, 2017.

[68] ——, “Cramer-Rao bound for indoor visible light positioning using an aperture-based angular-diversity receiver,” in *2016 IEEE Int. Conf. Commun. ICC 2016*, 2016, pp. 1–6.

[69] T. Q. Wang, Y. A. Sekercioglu, A. Neild, and J. Armstrong, “Position accuracy of time-of-arrival based ranging using visible light with application in indoor localization systems,” *J. Light. Technol.*, vol. 31, no. 20, pp. 3302–3308, 2013.

[70] S. Y. Jung, S. Hann, and C. S. Park, “TDOA-based optical wireless indoor localization using LED ceiling lamps,” *IEEE Trans. Consum. Electron.*, vol. 57, no. 4, pp. 1592–1597, 2011.

[71] A. Arafa, S. Dalmiya, R. Klukas, and J. F. Holzman, "Angle-of-arrival reception for optical wireless location technology," *Opt. Express*, vol. 23, no. 6, p. 7755, 2015.

[72] M. Yasir and S. Ho, "Indoor position tracking using multiple optical receivers," *J. Light. Technol.*, vol. 34, no. 4, pp. 1166–1176, 2016.

[73] Y. Hou, Y. Xue, C. Chen, and S. Xiao, "A RSS/AOA based indoor positioning system with a single LED lamp," in *2015 Int. Conf. Wirel. Commun. Signal Process. WCSP 2015*, 2015, pp. 1–4.

[74] G. B. Prince and T. D. C. Little, "A two phase hybrid RSS/AoA algorithm for indoor device localization using visible light," in *Globecom - Symp. Sel. Areas Commun.*, 2012, pp. 3347–3352.

[75] M. T. Taylor and S. Hranilovic, "Angular diversity approach to indoor positioning using visible light," in *2013 IEEE Globecom Work. (GC Wkshps)*. IEEE, 2013, pp. 1093–1098.

[76] W. Wang, J. Zhang, Q. Wang, and M. Zuniga, "Leveraging smart lights for passive localization," in *Proc. - 15th IEEE Int. Conf. Mob. Ad Hoc Sens. Syst. MASS 2018*, no. Section IV, 2018, pp. 272–280.

[77] W. Wang, Q. Wang, J. Zhang, and M. Zuniga, "PassiveVLP: Leveraging smart lights for passive positioning," *ACM Trans. Internet Things*, vol. 1, no. 1, pp. 1–25, 2019.

[78] B. Zhang, K. Ren, G. Xing, X. Fu, and C. Wang, “SBVLC: Secure barcode-based visible light communication for smartphones,” *IEEE Trans. Mob. Comput.*, vol. 15, no. 2, pp. 432–446, 2016.

[79] H. Hosseinianfar and M. Brandt-Pearce, “Cooperative object detection and parameter estimation using visible light communications,” 2020. [Online]. Available: http://arxiv.org/abs/2003.07525

[80] ——, “Performance limits for fingerprinting-based indoor optical communication positioning systems exploiting multipath reflections,” *IEEE Photonics J.*, vol. 12, no. 4, 2020.

[81] N. Faulkner, F. Alam, M. Legg, and S. Demidenko, “Watchers on the wall: Passive visible light-based positioning and tracking with embedded light-sensors on the wall,” *IEEE Trans. Instrum. Meas.*, vol. 69, no. 5, pp. 2522–2532, 2020.

[82] T. Li, C. An, Z. Tian, A. T. Campbell, and X. Zhou, “Human sensing using visible light communication,” *Proc. Annu. Int. Conf. Mob. Comput. Networking, MOBICOM*, vol. 2015-Septe, pp. 331–344, 2015.

[83] V. Nguyen, M. Ibrahim, S. Rupavatharam, M. Jawahar, M. Gruteser, and R. Howard, “Eyelight: Light-and-shadow-based occupancy estimation and room activity recognition,” *Proc. - IEEE INFOCOM*, pp. 351–359, 2018.

[84] M. Ibrahim, V. Nguyen, S. Rupavatharam, M. Jawahar, M. Gruteser, and R. Howard, “Visible light based activity sensing using ceiling photosensors,” in *Proc. Annu. Int. Conf. Mob. Comput. Networking, MOBICOM*, 2016, pp. 43–48.

[85] E. Di Lascio, A. Varshney, T. Voigt, and C. Perez-Penichet, "Poster abstract: LocaLight - A battery-free passive localization system using visible light," in *2016 15th ACM/IEEE Int. Conf. Inf. Process. Sens. Networks, IPSN 2016 - Proc.*, vol. 2, 2016, pp. 22–23.

[86] Z. Yang, Z. Wang, J. Zhang, C. Huang, and Q. Zhang, "Wearables can afford: Light-weight indoor positioning with visible light," in *Proc. 13th Annu. Int. Conf. Mob. Syst. Appl. Serv.* Florence: Association for Computing Machinery, 2015, pp. 317–330.

[87] K. Majeed and S. Hranilovic, "Passive indoor localization for visible light communication systems," in *2018 IEEE Glob. Commun. Conf.*, Abu Dhabi, 2018, pp. 1–6.

[88] ——, "Performance bounds on passive indoor positioning using visible light," *J. Light. Technol.*, vol. 38, no. 8, pp. 2190–2200, 2020.

[89] ——, "Passive indoor visible light positioning system using deep learning," *IEEE Internet Things J.*, 2021.

[90] ——, "Passive indoor visible light-based fall detection using neural networks," *under Prep.*, 2021.

[91] ——, "Passive positioning using visible light systems," in *Opt. Netw. Commun. Conf. & Exhib. 2021, OSA*, 2021.

[92] V. Pohl, V. Jungnickel, and C. Helmolt, "Integrating-sphere diffuser for wireless infrared communication," in *IEE Proc. - Optoelectron.*, vol. 147, no. 4, 2000, pp. 281–285.

[93] K. Lee, H. Park, and J. R. Barry, “Indoor channel characteristics for visible light communications,” *IEEE Commun. Lett.*, vol. 15, no. 2, pp. 217–219, 2011.

[94] A. M. Baldridge, S. J. Hook, C. I. Grove, and G. Rivera, “The ASTER spectral library version 2.0,” *Remote Sens. Environ.*, vol. 113, no. 4, pp. 711–715, 2009. [Online]. Available: http://dx.doi.org/10.1016/j.rse.2008.11.007

[95] S. K. Meerdink, S. J. Hook, D. A. Roberts, and E. A. Abbott, “The ECOSTRESS spectral library version 1.0,” *Remote Sens. Environ.*, vol. 230, no. May, p. 111196, 2019. [Online]. Available: https://doi.org/10.1016/j.rse.2019.05.015

[96] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning.* MIT Press, 2016. [Online]. Available: http://www.deeplearningbook.org

[97] A. M. Vegni and M. Biagi, “An indoor localization algorithm in a small-cell LED-based lighting system,” *2012 Int. Conf. Indoor Position. Indoor Navig. IPIN 2012 - Conf. Proc.*, pp. 1–7, 2012.

[98] M. Biagi, A. M. Vegni, and T. D. Little, “LAT indoor MIMO-VLC localize, access and transmit,” in *2012 Int. Work. Opt. Wirel. Commun. IWOW 2012*, no. 1, 2012, pp. 1–3.

[99] S. H. Madoka Nakajima, “Indoor navigation system for visually impaired people using visible light communication and compensated geomagnetic sensing,” in *IEEE Int. Conf. Commun. China Wirel. Commun. Syst.*, 2012, pp. 524–529.

[100] M. Nakajima and S. Haruyama, “New indoor navigation system for visually impaired people using visible light communication,” *EURASIP J. Wirel. Commun. Netw.*, vol. 2013, no. 1, p. 37, 2013.

[101] Z. Zhou, M. Kavehrad, and P. Deng, “Indoor positioning algorithm using light-emitting diode visible light communications,” *Opt. Eng.*, vol. 51, no. 8, aug 2012.

[102] A. Taparugssanagorn, S. Siwamogsatham, and C. Pomalaza-Raez, “A hexagonal coverage LED-ID indoor positioning based on TDOA with extended kalman filter,” in *Proc. - Int. Comput. Softw. Appl. Conf.*, 2013, pp. 742–747.

[103] J. Rudander, T. Husoy, and P. V. Walree, “Shallow-water channel sounding for high speed acoustic communication,” in *Ocean. 2017 - Aberdeen*, 2017, pp. 1–8.

[104] V. Jungnickel, V. Pohl, S. Nönnig, and C. von Helmolt, “A physical model of the wireless infrared communication channel,” *IEEE J. Sel. Areas Commun.*, vol. 20, no. 3, pp. 631–640, 2002.

[105] R. Schmidt, “Multiple emitter location and signal parameter estimation,” *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, 1986.

[106] J. R. Barry, J. M. Kahn, W. J. Krause, E. A. Lee, and D. G. Messerschmitt, “Simulation of multipath impulse response for indoor wireless optical channels,” *IEEE J. Sel. Areas Commun.*, vol. 11, no. 3, pp. 367–379, 1993.

[107] Q. Xu, R. Zheng, and S. Hranilovic, “IDyLL: Indoor localization using inertial and light sensors on smartphones,” in *Proc. ACM Int. Jt. Conf. Pervasive Ubiquitous Comput. UbiComp 2015*, no. SEPTEMBER 2015, 2015, pp. 307–318.

[108] Y. Kuo, P. Pannuto, K. Hsiao, P. Dutta, and A. Arbor, "Luxapose: Indoor positioning with mobile phones and visible light," in *Mobicom'14*, 2014, pp. 299–301.

[109] X. Liu, H. Makino, S. Kobayashi, and Y. Maeda, "Design of an indoor self-positioning system for the visually impaired - simulation with RFID and bluetooth in a visible light communication system," in *Annu. Int. Conf. IEEE Eng. Med. Biol. - Proc.*, vol. 4, no. 3, 2007, pp. 1655–1658.

[110] H. S. Kim, D. R. Kim, S. H. Yang, Y. H. Son, and S. K. Han, "An indoor visible light communication positioning system using a RF carrier allocation technique," *J. Light. Technol.*, vol. 31, no. 1, pp. 134–144, 2013.

[111] T. H. Do, J. Hwang, and M. Yoo, "TDoA based indoor visible light positioning systems," in *Int. Conf. Ubiquitous Futur. Networks, ICUFN*, 2013, pp. 456–458.

[112] U. Nadeem, N. Hassan, M. Pasha, and C. Yuen, "Highly accurate 3D wireless indoor positioning system using white LED lights," *Electron. Lett.*, vol. 50, no. 11, pp. 828–830, 2014.

[113] C. Sertthin, E. Tsuji, M. Nakagawa, S. Kuwano, and K. Watanabe, "A switching estimated receiver position scheme for visible light based indoor positioning system," in *2009 4th Int. Symp. Wirel. Pervasive Comput.*, 2009, pp. 1–5.

[114] G. Del Campo-Jimenez, J. M. Perandones, and F. J. Lopez-Hernandez, "A VLC-based beacon location system for mobile applications," in *2013 Int. Conf. Localization GNSS, ICL-GNSS 2013*, no. Id, 2013, pp. 3–6.

[115] S. Han, Y. Won, D. Kim, and S. Yang, “Three-dimensional optical wireless indoor positioning system using location code map based on power distribution of visible light emitting diode,” *IET Optoelectron.*, vol. 7, no. 3, pp. 77–83, 2013.

[116] P. Luo, M. Zhang, X. Zhang, G. Cai, D. Han, and Q. Li, “An indoor visible light communication positioning system using dual-tone multi-frequency technique,” in *Int. Work. Opt. Wirel. Commun.*, 2013, pp. 25–29.

[117] S. Han, E. Jeong, D. Kim, Y. Son, S. Yang, and H. Kim, “Indoor three-dimensional location estimation based on LED visible light communication,” *Electron. Lett.*, vol. 49, no. 1, pp. 54–56, 2013.

[118] S. Jung, S. R. Lee, and C. Park, “Indoor location awareness based on received signal strength ratio and time division multiplexing using light-emitting diode light,” *Opt. Eng.*, vol. 53, pp. 1–6, 2014.

[119] W. Gu, M. Aminikashani, P. Deng, and M. Kavehrad, “Impact of multipath reflections on the performance of indoor visible light positioning systems,” *J. Light. Technol.*, vol. 34, no. 10, pp. 2578–2587, 2016.

[120] K. Majeed and S. Hranilovic, “Passive indoor localization for visible light communication systems,” in *2018 IEEE Glob. Commun. Conf.*, Abu Dhabi, 2018, pp. 1–6.

[121] J. H. Y. Nah, R. Parthiban, and M. H. Jaward, “Visible light communications localization using TDOA-based coherent heterodyne detection,” in *4th Int. Conf. Photonics, ICP 2013 - Conf. Proceeding*, 2013, pp. 247–249.

[122] S. Lee and S. Y. Jung, "Location awareness using angle-of-arrival based circular-PD-array for visible light communication," in *APCC 2012 - 18th Asia-Pacific Conf. Commun. "Green Smart Commun. IT Innov.*, 2012, pp. 480–485.

[123] M. F. Keskin and S. Gezici, "Comparative theoretical analysis of distance estimation in visible light positioning systems," *J. Light. Technol.*, vol. 34, no. 3, pp. 854–865, 2016.

[124] M. F. Keskin, S. Gezici, and O. Arikan, "Direct positioning in synchronous and asynchronous visible light systems," *IEEE Trans. Commun.*, vol. 66, no. 1, pp. 239–254, 2018.

[125] E. Gonendik and S. Gezici, "Fundamental limits on RSS based range estimation in visible light positioning systems," *IEEE Commun. Lett.*, vol. 19, no. 12, pp. 2138–2141, 2015.

[126] C. Amini, A. Taherpour, T. Khattab, and S. Gazor, "Theoretical accuracy analysis of indoor visible light communication positioning system based on time-of-arrival," in *Can. Conf. Electr. Comput. Eng.*, 2016, pp. 1–5.

[127] N. Stevens and H. Steendam, "Magnitude of the distance estimation bias in received signal strength visible light positioning," *IEEE Commun. Lett.*, vol. 22, no. 11, pp. 2250–2253, 2018.

[128] H. L. V. Trees, *Detection, Estimation, and Modulation Theory, Part I.* John Wiley & Sons, Inc., 2001.

[129] T. Do and M. Yoo, "TDOA-based indoor positioning using visible light," *Photonic Netw. Commun.*, vol. 27, no. 2, pp. 80–88, 2014.

[130] M. S. Davis, G. A. Showman, and A. D. Lanterman, "Coherent MIMO radar: The phased array and orthogonal waveforms," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 29, no. 8, pp. 76–91, 2014.

[131] J. Grubor, S. Randel, K. D. Langer, and J. W. Walewski, "Broadband information broadcasting using LED-based interior lighting," *J. Light. Technol.*, vol. 26, no. 24, pp. 3883–3892, 2008.

[132] Zemax OpticStudio, "Optical system design software," 2020. [Online]. Available: https://www.zemax.com/

[133] P. Dutre, E. Lafortune, and Y. D. Willems, "A mathematical framework for global illumination algorithms," in *Proc. Winter Sch. Comput. Graph. CAD Syst. '94*, 1994, pp. 75–84.

[134] S. Rodríguez Pérez, R. Pérez Jiménez, F. J. López Hernández, O. B. González Hernández, and A. J. Ayala Alfonso, "Reflection model for calculation of the impulse response on IR-wireless indoor channels using ray-tracing algorithm," *Microw. Opt. Technol. Lett.*, vol. 32, no. 4, pp. 296–300, 2002.

[135] H. B. Eldeeb, M. Uysal, S. M. Mana, P. Hellwig, J. Hilt, and V. Jungnickel, "Channel modelling for light communications: Validation of ray tracing by measurements," in *2020 12th Int. Symp. Commun. Syst. Networks Digit. Signal Process.*, Porto, 2020, pp. 1–6.

[136] D. Li, C. Gong, and Z. Xu, "A RSSI-based indoor visible light positioning approach," in *2016 10th Int. Symp. Commun. Syst. Networks Digit. Signal Process. CSNDSP 2016*, 2016, pp. 1–6.

[137] W. Zhang and M. Kavehrad, “A 2-D indoor localization system based on visible light LED,” in *2012 IEEE Photonics Soc. Summer Top. Meet. Ser. PSST 2012*, vol. 4, 2012, pp. 80–81.

[138] S. H. Yang, H. S. Kim, Y. H. Son, and S. K. Han, “Three-dimensional visible light indoor localization using AOA and RSS with multiple optical receivers,” *J. Light. Technol.*, vol. 32, no. 14, pp. 2480–2485, 2014.

[139] Y. S. Erogluy, I. Guvency, N. Palay, and M. Yukselz, “AOA-based localization and tracking in multi-element VLC systems,” in *2015 IEEE 16th Annu. Wirel. Microw. Technol. Conf. WAMICON 2015*, 2015.

[140] S. Han, E. Jeong, S. Yang, and H. Kim, “Tilted receiver angle error compensated indoor positioning system based on visible light communication,” *Electron. Lett.*, vol. 49, no. 14, pp. 890–892, 2013.

[141] Y. Kim, “VLC-TDOA using sinusodial pilot signal,” in *Int. Conf. IT Converg. Secur.*, Macao, 2013, pp. 1–3.

[142] M. Pasha, C. Yuen, N. Hassan, and U. Nadeem, “Indoor positioning system designs using visible LED lights: Performance comparison of TDM and FDM protocols,” *Electron. Lett.*, vol. 51, no. 1, pp. 72–74, 2015.

[143] S. Yang, D. Kim, H. Kim, Y. Son, and S. Han, “Indoor positioning system based on visible light using location code,” in *Fourth Int. Conf. Commun. Electron.*, 2012, pp. 360–363.

[144] S. Yang, E. Jung, and S. Han, "Indoor location estimation based on LED visible light communication using multiple optical receivers," *IEEE Commun. Lett.*, vol. 17, no. 9, pp. 1834–1837, 2013.

[145] W. Zhang, M. I. S. Chowdhury, and M. Kavehrad, "Asynchronous indoor positioning system based on visible light communications," *Opt. Eng.*, vol. 53, no. 4, p. 045105, 2014.

[146] M. Yasir, S. Ho, and B. N. Vellambi, "Indoor positioning system using visible light and accelerometer," *J. Light. Technol.*, vol. 32, no. 19, pp. 3306–3316, 2014.

[147] D. Konings, N. Faulkner, F. Alam, E. M. Lai, and S. Demidenko, "FieldLight: Device-free indoor human localization using passive visible light positioning and artificial potential fields," *IEEE Sens. J.*, vol. 20, no. 2, pp. 1054–1066, 2020.

[148] H. Hosseinianfar and M. Brandt-Pearce, "Cooperative passive pedestrian detection and localization using a visible light communication access network," *IEEE Open J. Commun. Soc.*, vol. 1, no. August, pp. 1325–1335, 2020.

[149] F. R. Gfeller and U. Bapst, "Wireless in-house data communication via diffuse infrared radiation," *IEEE Proc.*, vol. 67, no. 1, pp. 1474–1486, 1979.

[150] Blender, "Open source 3D creation software," 2020. [Online]. Available: https://www.blender.org/

[151] J. J. Goldberger and J. Ng, *Practical Signal and Image Processing in Clinical Cardiology*, 1st ed. Springer, London, 2010.

[152] E. Sarbazi, M. Uysal, M. Abdallah, and K. Qaraqe, "Ray tracing based channel modeling for visible light communications," in *2014 22nd Signal Process. Commun. Appl. Conf.* Trabzon: IEEE, 2014, pp. 702–705.

[153] F. Miramirkhani and M. Uysal, "Channel modeling and characterization for visible light communications," *IEEE Photonics J.*, vol. 7, no. 6, pp. 1–16, 2015.

[154] A. Al-Kinani, C. X. Wang, L. Zhou, and W. Zhang, "Optical wireless communication channel measurements and models," *IEEE Commun. Surv. Tutorials*, vol. 20, no. 3, pp. 1939–1962, 2018.

[155] F. Miramirkhani and M. Uysal, "Channel modelling for indoor visible light communications," *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.*, vol. 378, 2020.

[156] R. H. Hahnloser, R. Sarpeshkar, M. A. Mahowald, R. J. Douglas, and H. S. Seung, "Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit," *Nature*, vol. 408, no. 6815, pp. 947–951, 2000.

[157] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[158] T. Dietterich, "Overfitting and undercomputing in machine learning," *ACM Comput. Surv.*, vol. 27, no. 3, pp. 326–327, 1995.

[159] The MathWorks Inc., "Deep learning toolbox," Natick, Massachusetts, United States, 2020. [Online]. Available: https://www.mathworks.com/help/deeplearning/

[160] R. W. Broadley, J. Klenk, S. B. Thies, L. P. Kenney, and M. H. Granat, "Methods for the real-world evaluation of fall detection technology: A scoping review," *Sensors (Basel).*, vol. 18, no. 7, pp. 1–28, 2018.

[161] X. Wang, J. Ellul, and G. Azzopardi, "Elderly fall detection systems: A literature survey," *Front. Robot. AI*, vol. 7, no. June, 2020.

[162] Y. Nizam, M. N. H. Mohd, and M. M. A. Jamil, "Human fall detection from depth images using position and velocity of subject," *Procedia Comput. Sci.*, vol. 105, no. Iris 2016, pp. 131–137, 2017.

[163] X. Wang and K. Jia, "Human fall detection algorithm based on YOLOv3," *2020 IEEE 5th Int. Conf. Image, Vis. Comput. ICIVC 2020*, pp. 50–54, 2020.

[164] J. Zhang, C. Wu, and Y. Wang, "Human fall detection based on body posture spatio-temporal evolution," *Sensors (Switzerland)*, vol. 20, no. 3, 2020.

[165] Q. Guan, C. Li, X. Guo, and B. Shen, "Infrared signal based elderly fall detection for in-home monitoring," *Proc. - 9th Int. Conf. Intell. Human-Machine Syst. Cybern. IHMSC 2017*, vol. 1, pp. 373–376, 2017.

[166] Y. T. Chang and T. K. Shih, "Human fall detection based on event pattern matching with ultrasonic array sensors," *Ubi-Media 2017 - Proc. 10th Int. Conf. Ubi-Media Comput. Work. with 4th Int. Work. Adv. E-Learning 1st Int. Work. Multimed. IoT Networks, Syst. Appl.*, pp. 8–11, 2017.

[167] M. Gietzelt, J. Spehr, Y. Ehmen, S. Wegel, F. Feldwieser, M. Meis, M. Marschollek, K. Wolf, E. Steinhagen-Thiessen, and M. Gövercin,

"GAL@Home: A feasibility study of sensor-based in-home fall detection," *Z. Gerontol. Geriatr.*, vol. 45, pp. 716–721, dec 2012.

[168] O. Aziz, M. Musngi, E. J. Park, G. Mori, and S. N. Robinovitch, "A comparison of accuracy of fall detection algorithms (threshold-based vs. machine learning) using waist-mounted tri-axial accelerometer signals from a comprehensive set of falls and non-fall trials," *Med. Biol. Eng. Comput.*, vol. 55, no. 1, pp. 45–55, 2017.

[169] S. Palipana, D. Rojas, P. Agrawal, and D. Pesch, "FallDeFi: Ubiquitous fall detection using commodity Wi-Fi devices," *Proc. ACM Interactive, Mobile, Wearable Ubiquitous Technol.*, vol. 1, no. 4, pp. 1–25, 2018.

[170] Y. Huang, W. Chen, H. Chen, L. Wang, and K. Wu, "G-Fall: Device-free and training-free fall detection with geophones," in *2019 16th Annu. IEEE Int. Conf. Sensing, Commun. Netw.* IEEE, 2019, pp. 1–9.