

Visible Light Positioning Based Robotic Navigation and Mapping

by

Yiru WANG

A Thesis Submitted to
The Hong Kong University of Science and Technology
in Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy
in the Department of Electronic and Computer Engineering

April 2022, Hong Kong

Authorization

I hereby declare that I am the sole author of the thesis.

I authorize the Hong Kong University of Science and Technology to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize the Hong Kong University of Science and Technology to reproduce the thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

Signature redacted

Yiru WANG

April 2022

Visible Light Positioning Based Robotic Navigation and Mapping

by

Yiru WANG

This is to certify that I have examined the above PhD thesis
and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by
the thesis examination committee have been made.

Signature redacted

Prof. C. Patrick YUE, ECE Department (Thesis Supervisor)

Signature redacted

Prof. Andrew POON (Head of ECE Department)

Thesis Examination Committee:

1. **Prof. C. Patrick YUE** (Supervisor), Department of Electronic and Computer Engineering
2. **Prof. Ling SHI**, Department of Electronic and Computer Engineering
3. **Prof. Shaojie SHEN**, Department of Electronic and Computer Engineering
4. **Prof. Chi Keung TANG**, Department of Computer Science and Engineering
5. **Prof. Xiaolin TANG** (External Examiner), Department of Mechanical and Vehicle Engineering, Chongqing University

Department of Electronic and Computer Engineering,
The Hong Kong University of Science and Technology

April 2022

To my family

Acknowledgements

First of all, I would like to express my deepest gratitude to my supervisor, Prof. C. Patrick Yue for his patience, valuable guidance and strong support along the way. I have benefited greatly from his immense knowledge and plentiful experience, which encouraged me in my academic research as well as my daily life. It has been a wonderful experience to pursue my Ph.D. degree in Prof. Yue's group. I have not only learned how to conduct research and keep motivated, but also learned how to be happy and strong in life.

In addition, I would like to thank Prof. Ling Shi, Prof. Shaojie Shen, Prof. Chi Keung Tang, and Prof. Xiaolin Tang for taking time out of their busy schedules to serve as my thesis committee members and propose invaluable comments and insightful feedback to my thesis and research.

I would also like to offer my special thanks to my colleagues, both in and outside of my research group, including Dr. Hussain Babar, Dr. Wang Li, Dr. Weimin Shi, Dr. Jeff Qiu, Dr. Xianbo Li, Dr. Kuang Fang, Mr. Sam Wang, Mr. Jian Kang, Ms. Xuan Wu, Ms. Bo Xu, Mr. Fredrick Hong, Mr. Rehan Azmat, Ms. Zilu Liu, Ms. Tianxin Min, Ms. Xinyi Liu, Mr. Chongyun Zhang, Mr. Fuzhan Chen, Mr. Weishun Deng, Mr. Johnny Cheng, Mr. Hamed Fallah, and Mr. Johar Abdekhoda, all of whom were ready to help and shared numerous valuable discussions with me whenever needed. They were wonderful lab mates as well as good friends. I should also acknowledge the technical and logistics support from the ECE lab technician, Mr. Raymond Cheng.

I am thankful to my college friends, Ms. Hailin Zhou, Dr. Bowei Xu, Mr. Junjie Zhang and Mr. Zhili Su, who cheered me on and gave me support outside my Ph.D. studies. I would never forget the laughter and happiness we had together.

Finally, my sincere gratitude goes to my parents, Yunyan Gao and Chen Wang, for their unwavering support and endless love which keep me motivated and confident. Without them, I would not have made it through my Ph.D. degree.

Table of Contents

Authorization Page	ii
CHAPTER 1 Introduction to VLC Systems.....	1
1.1 Introduction to VLC.....	1
1.2 Types of VLC Systems Based on a Receiver	2
1.2.1 High-Data-Rate PD-based VLC Systems.....	2
1.2.2 Low-Data-Rate Image Sensor-based VLC Systems.....	3
1.3 Applications of VLC Systems.....	5
1.3.1 Indoor VLC Systems	5
1.3.2 Outdoor VLC Systems	7
1.4 Related Works.....	8
1.5 Thesis Organization	9
1.6 References.....	9
CHAPTER 2 Visible Light Positioning Systems	13
2.1 Positioning Technologies.....	13
2.2 VLP System Configuration.....	15
2.2.1 PD-based VLP Systems.....	15
2.2.2 Image Sensor-based VLP Systems.....	17
2.3 VLP Technology Applications.....	19
2.3.1 Location-aware Services	19
2.3.2 Robotics Navigation and Localization	19
2.4 References.....	21
CHAPTER 3 High-Accuracy Indoor VLP Systems for Mobile Devices.....	25
3.1 Introduction.....	25
3.2 Related Works.....	26
3.3 Arbitrarily Tilted Receiver Camera Correction Method for VLP Systems	28
3.3.1 Camera Model and Imaging Geometry	29
3.3.2 Tilted Receiver Camera Correction Method	31
3.4 Partially Blocked Led Image Compensation Method for VLP Systems.....	35
3.5 Experiment and Evaluation.....	37
3.5.1 Performances of the Partially Blocked LED Image Compensation Methods	39
3.5.2 Performances of the Tilted Receiver Camera Correction and Partially Blocked LED Image Compensation Methods for a Single-LED VLP System.....	40
3.6 Summary	44

3.7 References.....	44
CHAPTER 4 High-Accuracy Indoor VLP System for Robots	48
4.1 Indoor Positioning for Robots.....	48
4.2 Related Works.....	49
4.2.1 Position Estimation of Robot	49
4.2.2 Path Planning Algorithms	51
4.3 VLP-based Mobile Robot Platform	53
4.3.1 Smart LED System for VLP.....	54
4.3.2 Real-time Robotic Localization System.....	56
4.4 Application of Robot Positioning and Navigation System: Panorama Creation	59
4.4.1 Real-time Robotic Navigation System	59
4.4.2 Image Taking at a Target Point	60
4.4.3 Panorama Creation	65
4.5 Summary	66
4.6 References.....	66
CHAPTER 5 VLP and SLAM-assisted Map Calibration for Robot Navigation.....	70
5.1 Introduction.....	70
5.2 Related Work	72
5.3 Mapping System	74
5.3.1 Occupancy Grid Mapping System.....	74
5.3.2 Map Transformation.....	74
5.3.3 Overview of the Proposed Map Calibration System	77
5.4 Map Calibration Method.....	78
5.4.1 Positioning on Two Different Maps	78
5.4.2 Calibration of the Orientation.....	80
5.4.3 Calibration of the Scale	83
5.5 Experimental Results	86
5.5.1 Experiment Setup	86
5.5.2 Mapping Process and Alignment Results.....	87
5.5.3 Navigation on Calibrated Map	90
5.5.4 Navigation with Semantic Information	92
5.6 Summary	94
5.7 References.....	95

CHAPTER 6 Conclusion, Future Work and Publications.....	98
6.1 Summary of Contributions.....	98
6.2 Future Work.....	99
6.2.1 An Prior Information Assisted Distortion Elimination Method for Occupancy Grid Map Construction	99
6.2.2 Automated 3D Reconstruction using Robot-mounted 360-Degree Camera with Visible Light Positioning Technology for Building Information Modelling Applications	100
6.3 Publications.....	102
6.3.1 Works under review	102
6.3.2 Published	103

List of Figures

Figure 1.1: Spectrum of RF versus visible light.	1
Figure 1.2: A PD-based VLC system.	2
Figure 1.3: The method of image sensor scanning the image: (a) rolling shutter mode, (b) global shutter mode.....	3
Figure 1.4: An image sensor-based VLC system.	4
Figure 1.5: Transceivers of SCC systems.	5
Figure 1.6: An example for a 2D COBRA barcode.	6
Figure 2.1: An indoor VLP system based on a single PD and four LEDs.	16
Figure 2.2: Block diagram of a smart phone based VLC system illustrating (a) transmitter side and (b) receiver side [22].....	18
Figure 2.3: Diagram of a common indoor VLP system with multiple mobile robot collaboration and human-robot cooperation under based on ceiling mounted LED lightings.	20
Figure 3.1: The proposed tilted receiver camera correction method for a single-LED-based VLP system: (1) finding the LED's world coordinates, (2) computing the camera matrices.	29
Figure 3.2: Imaging geometry of a single-LED-based VLP system with a tilted camera.....	29
Figure 3.3: Rotation angles of a smartphone along x-, y- and z-axes.	31
Figure 3.4: Images captured at the same position with different rotation angles: (a) $\varphi_x = 0^\circ$, $\varphi_y = 0^\circ$; (b) $\varphi_x = 15^\circ$, $\varphi_y = -9^\circ$; (c) $\varphi_x = 17^\circ$, $\varphi_y = 3^\circ$; (d) $\varphi_x = -20^\circ$, $\varphi_y = -2^\circ$	32
Figure 3.5: The image processing procedures of boundary detection: (a) original image, (b) binary image, (c) patterns eliminated, (d) detected boundary, (e) detected center in red and major axis in green.	32
Figure 3.6: The pinhole camera model describing the relationship between the major axis length of the LED image and the distance between the LED and the camera.	34
Figure 3.7: Incomplete LED images caused by a limited FOV, blockage of a human head and furniture.	35
Figure 3.8: The detected shadow curve caused by human head blockage.	36
Figure 3.9: The experimental setup of the proposed single LED-based VLP system.	38

Figure 3.10: The detection results of center searching method and boundary fitting method: (a) the original captured LED image, (b) the detection results, (c) the zoomed in detection results.	39
Figure 3.11: The comparison of the positioning error performance based on LED images with different captured area ratios among the two proposed methods, the RSS-AOA method [16] and positioning without rotation correction.	40
Figure 3.12: The comparison of average PERs based on LED images with different captured area ratios among the two proposed methods and the circle geometry method [15], RSS-AOA method [16] and projective geometry method [17].	42
Figure 3.13: Coverage extension of the proposed center searching method and boundary fitting method along the margin of FOV.	43
Figure 4.1: Diagram illustrating VLC sources used for robotic 3D localization.	54
Figure 4.2: Demonstration setup of high-precision positioning system based on VLC smart lighting.	55
Figure 4.3: Experiment results of the proposed robot positioning system: (a) experiment setup, (b) VLP results when the camera moves on a plane, (c) VLP results when the camera rotates and is partially blocked in 3D space.	58
Figure 4.4: 360-degree image capture: (a) a 360-degree camera, (b) a typical 360-degree image.	59
Figure 4.5: The robot equipped with a typical USB camera in the proposed VLP-based image taking robot system.	61
Figure 4.6: Diagram of the proposed VLP-based image taking robot system.	61
Figure 4.7: Graph representation of the nodes in the proposed ROS-based image taking robot system.	62
Figure 4.8: Navigation and image saving process: (a) sending a navigation goal to the robot via RViz, (b) the robot moving to the target point, (c) the robot saving images after arriving, (d) sending another navigation goal to the robot via a command line.	64
Figure 4.9: The saved images using the proposed system: (a) the created folders named with the x- and y-coordinates estimated by VLP, (b) the saved images named with the azimuth angle.	65
Figure 4.10: The generated panorama using the saved images with location information.	66
Figure 5.1: Occupancy-coordinate transformation for a grayscale map image.	75

Figure 5.2: The proposed VLP landmarks and SLAM-assisted automatic map calibration for robotic navigation: (1) system setup, (2) mapping process, (3) calibration process.	77
Figure 5.3: Mapping process with VLP landmarks and SLAM technology.	79
Figure 5.4: Key points on the two maps: (a) layout map, (b) sensor map.	80
Figure 5.5: Orientation calibration for the sensor map: (a) translate the rotation center, (b) rotate and crop, (c) fill in the corners.....	81
Figure 5.6: Vertical and horizontal distances between the two key points: (a) on the layout map, (b) on the sensor map.....	83
Figure 5.7: Experimental setup: (a) a building blueprint, (b) a floor map, (c) three VLP landmarks, (d) a TurtleBot3 Burger robot.	86
Figure 5.8: Mapping process visualized in RViz.....	88
Figure 5.9: Map alignment result of the building blueprint and sensor map: (a) based on two key points, (b) based on three key points.	89
Figure 5.10: Navigation goal and navigation results on the sensor map, floor map and building blueprint.	90
Figure 5.11: Navigation process with LiDAR detection results, DWA local plan and Dijkstra's global plan.....	90
Figure 5.12: Compiled semantic information on the floor map.	92
Figure 5.13: Navigation process with semantic information: (a) sending the command to the robot, (b) robot navigating to the target point, (c) after reaching the target point, robot waiting for the next navigation goal, (d) robot navigating to the second target point.....	94
Figure 6.1: Conceptual diagram illustrating 3D reconstruction using 360-degree cameras mounted on multiple mobile robots, and visible light communication (VLC) sources used for 3D localization and camera pose estimation.	102

List of Tables

Table 2.1: Summary of RF-based IPS [15].	14
Table 2.2: Comparison of different PD-based VLP systems [15].	17
Table 2.3: Advantages of VLP in comparison to SLAM-based positioning systems.....	21
Table 3.1: Experimental Parameters	39
Table 4.1: Comparison between global path planning and local path planning.....	52
Table 5.1: Experimental Parameters	86
Table 5.2: Navigation Results	91

Visible Light Positioning Based Robotic Navigation and Mapping

by

Yiru WANG

Department of Electronic and Computer Engineering,
The Hong Kong University of Science and Technology

Abstract

With the growing demand for location-based services such as indoor navigation, robot control and object tracking, indoor positioning technology has attracted increasing attention from both academia and industry. For outdoor environments, the Global Positioning System (GPS) provides real-time positioning services based on satellites and is widely used in airplanes, automobiles and portable devices. However, it cannot realize efficient positioning in indoor environments because satellite signals will be extremely attenuated and interrupted by indoor obstacles. Currently, wireless technologies, including Bluetooth and WiFi, are widely applied to indoor positioning systems. However, these technologies can only achieve meter-level accuracy and are potentially vulnerable to malicious activities. Visible light positioning (VLP) technology can solve these problems, with multiple advantages including centimeter-level accuracy, compatibility with existing lighting infrastructure, low cost and insusceptibility to electromagnetic interference. Therefore, VLP systems are very competitive to provide indoor positioning service. In this thesis, a high-accuracy VLP system is proposed, based on which robotic navigation and map construction is also achieved. The design and implementation of the system is divided into three parts.

In the first part, an image sensor-based single-LED VLP system is proposed. The additional positioning error caused by tilted receiver camera is corrected by the rotation angles estimated

by the inertial sensors. The proposed VLP system can also provide positioning services even when an incomplete LED image is captured by the camera.

In the second part, a VLP-based mobile robot experiment platform is built. The proposed platform consists of two parts: intelligent lighting and image sensor-based VLP light tracking. Smart LEDs are used as the access points of the VLP system and are modulated with digital IDs containing the information of the LEDs' world coordinates. Therefore, the proposed positioning system is scalable, with no maximum scale limit. The camera mounted on robot will capture the images of LEDs and use ID recognition algorithm to identify the IDs then get the position with geometric feature-based image processing algorithm. Based on the proposed robot positioning and navigation system, a panorama creation method is proposed which can generate a panorama at any target point using a robot mounted with an ordinary USB camera.

In the final part, an autonomous map construction method using VLP landmarks and Simultaneous Localization and Mapping (SLAM). A layout map of the environment to be perceived is calibrated by a robot tracking at least two landmarks mounted in the venue. At the same time, the robot's position on the occupancy grid map generated by SLAM is recorded. A map transformation method is then performed to align the orientation of the two maps and to calibrate the scale of the layout map to agree with that of the sensor map. After the calibration, the semantic information on the layout map remains and the accuracy is improved.

CHAPTER 1 Introduction to VLC Systems

1.1 Introduction to VLC

Due to the rapid growth of the Internet of Things (IoT) and increasing demand for wireless services, it has become more and more difficult for the limited spectrum resources of radio frequency (RF) systems to meet the needs of wireless users. For visible light spectrum, about 390 THz of bandwidth is available, while the entire RF spectrum, including microwaves, consists of 300 GHz of bandwidth [1], as shown in Figure 1.1. Therefore, the visible light spectrum is about 1300 times the size of the entire RF spectrum [2]. Visible light communication (VLC) systems utilizing unlicensed light spectrum can serve as an alternative technology to the existing RF systems in indoor wireless applications.

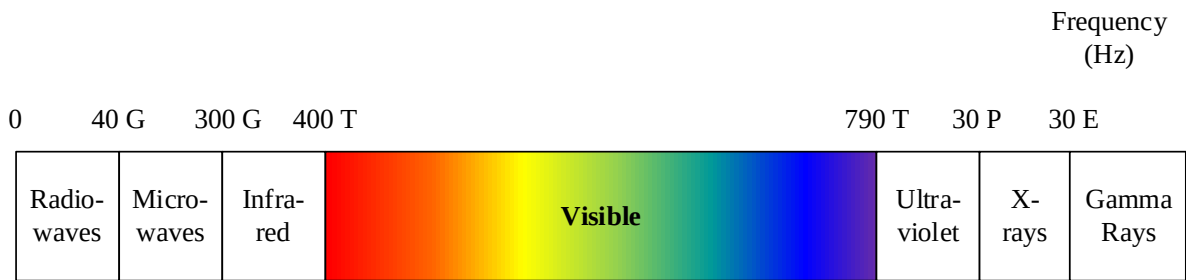


Figure 1.1: Spectrum of RF versus visible light.

Different from earlier lighting facilities, LEDs are capable of achieving high-speed response to light intensity modulation (IM). By encoding data in the emitted light, LEDs can be utilized as access points in VLC systems to transmit information at a high speed [3]. If the modulation frequency is higher than the flicker fusion threshold [4], the human eyes will not observe the changing light intensity levels, and LEDs can transmit data and maintain the illumination quality at the same time. Furthermore, LEDs can provide long-lifetime, low-power, high-brightness and stable illumination services with a wide band light beam, leading to its increasing demand of employment for general lighting in commercial and residential scenarios [3]. Therefore, VLC systems have multiple advantages by using LEDs as transmitters, including low cost, high signal-to-noise ratio (SNR), anti-disturbance of electromagnetism and good confidentiality. As visible light beam cannot penetrate through the walls and most non-

transparent objects in buildings, which is known as the light-of-sight (LOS) property of visible light beam, VLC signals in different rooms will not interference with each other, and therefore are independent and private.

1.2 Types of VLC Systems Based on a Receiver

The receivers in VLC systems can be divided into two types: photodiodes (PDs) and image sensors. Compared with PD-based VLC, image sensor-based VLC, which is also known as optical camera communication (OCC) [5], is more attractive due to the ubiquity of cameras on most of today's mobile devices, such as smartphones and tablets.

1.2.1 High-Data-Rate PD-based VLC Systems

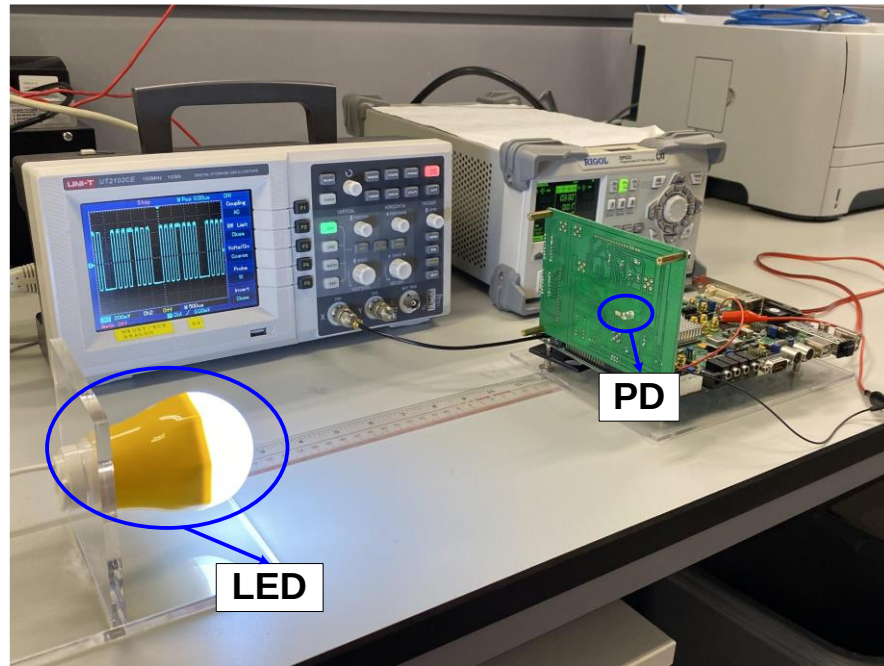


Figure 1.2: A PD-based VLC system.

In a PD-based VLC system as shown in Figure 1.2, the LED-based transmitter is controlled by a VLC-enabled LED driver to transmit optical waveforms and the PD-based receiver converts the received optical power to electrical current following the photoelectric effect. Compared with an image sensor-based VLC system, a PD-based system can achieve higher data rate. Stand-alone photodetectors are capable to achieve throughputs of several gigabits per second [6]. However, the detection areas of PDs are very small. Therefore, PD-based VLC

systems require precise alignment between LED and PD for signal detection [7], which limits its application in our daily lives.

1.2.2 Low-Data-Rate Image Sensor-based VLC Systems

When optical waves falls on an image sensor, they are collected by an array of small discrete cells called pixels. The projected image can thereby be divided into these pixels. The quantum efficiency of a pixel determines its efficiency of absorbing incident photons and converting incident photons into charges. Each pixel consists of a PD and an image sensor is composed of a matrix of multiple PDs.

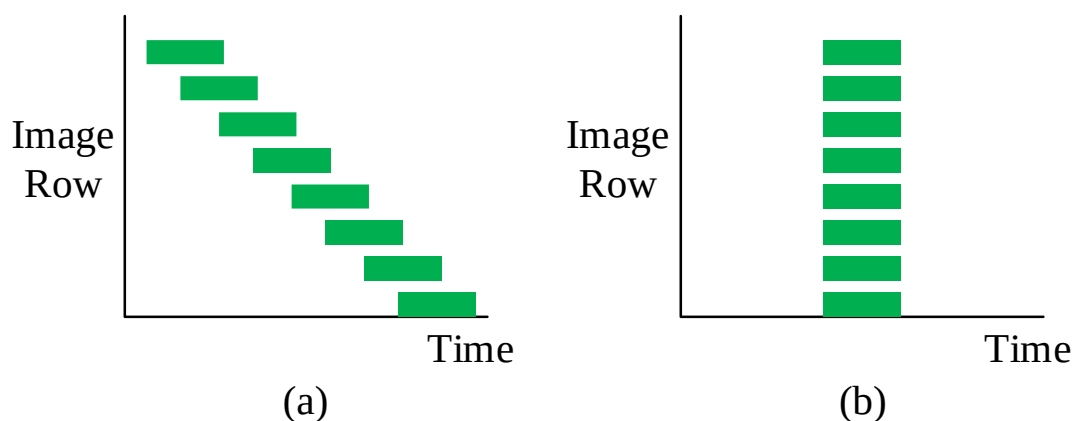


Figure 1.3: The method of image sensor scanning the image: (a) rolling shutter mode, (b) global shutter mode.

There are two major methods of image sensors reading out the signal from each pixel, being the global shutter mode and the rolling shutter mode, as shown in Figure 1.3. An image sensor working at global shutter mode capture an entire frame all at once. All the pixels on the sensor are read out at the same time. Most charge-coupled device (CCD) sensors employ global shutter mode. The “rolling shutter” feature of an image sensor can be used to receive data at a faster rate. Since there are a huge number of PDs in one image sensor, it is not possible to read out all the pixels at the same time. Therefore, at a time, only one row or one column in the PD matrix is read out, where the lines are scanned sequentially. This process of reading the output of each pixel row by row or column by column is called rolling shutter. Complementary metal–oxide–semiconductor (CMOS) sensors tend to work at rolling shutter mode. Due to the huge

prevalence of CMOS sensor in modern cameras, rolling shutter has been extensively used in video and filmmaking. One main drawback of rolling shutter is that it will cause warping when capturing a moving object, because the camera working in rolling shutter mode is always exposed in a progressive motion.

It is noteworthy that the CMOS image sensor on any mobile device can receive signals via VLC. In order to achieve a high resolution, a large amount of PDs have to be placed on one image sensor. A modern image sensor usually contains up to millions of discrete PDs, which leads to a low frame rate that the image sensor can achieve. For example, the frame rate of a conventional image sensor on a smartphone is usually lower than 40 fps [8].

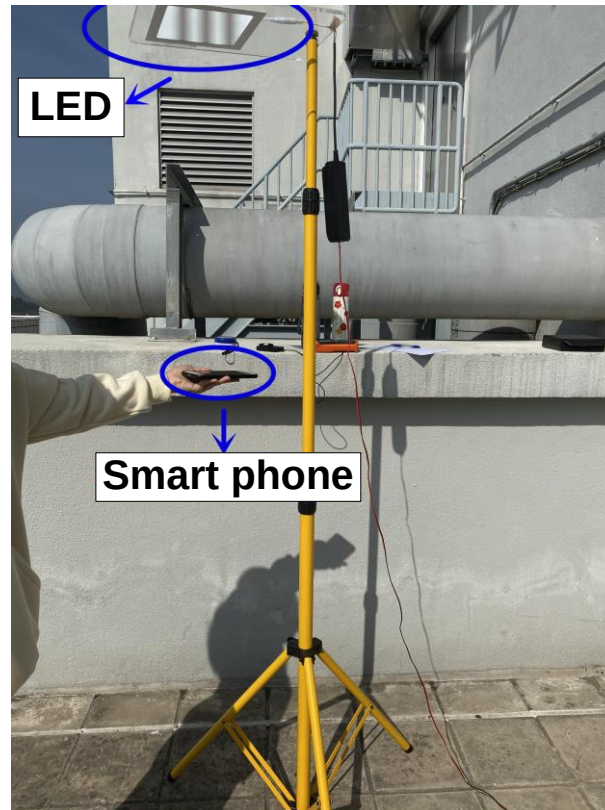


Figure 1.4: An image sensor-based VLC system.

Figure 1.4 shows a VLC system using an image sensor working at rolling shutter mode. The bright bars captured by the image sensor correspond to the transmitted data 1 and the dark bars correspond to the transmitted data 0. Then at the receiver side, the CMOS image sensor on the smart phone captures the pixel from top to bottom [7], extracts the rolling shutter patterns by image processing and decodes the rolling shutter patterns.

1.3 Applications of VLC Systems

According to application scenarios, VLC technology can be divided into indoor and outdoor applications. Driven by the success of the Li-Fi concept [9], indoor VLC has attracted more interest and achieved higher growth. However, due to more crucial environments and strict constraints, such as regulation, mobility and weather, outdoor VLC has been developed at a lower speed, but still have achieved significant results.

1.3.1 Indoor VLC Systems

As a part of image sensor-based VLC, screen to camera communication (SCC) can provide high-reliability and high-flexibility short-range communication services, and therefore has gained much interest from both industry and academia [10]. SCC systems encode data bits into images and then display the images on a liquid crystal display (LCD) screen, such as a smartphone or laptop screen, as shown in Figure 1.5. At the receiver side, the user will use a camera to record the images and then decode the data blocks.

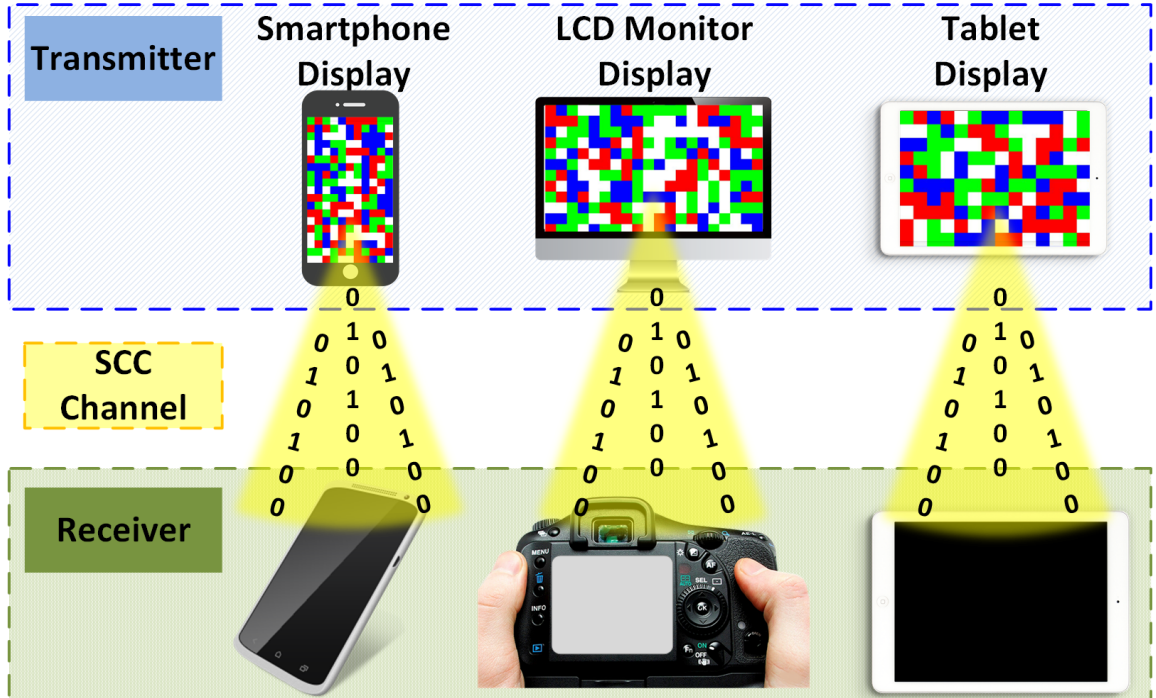


Figure 1.5: Transceivers of SCC systems.

An SCC system based on 2D color barcode, called COBRA: Color Barcode stReaming for sMARTphones, is proposed in [10]. In a COBRA system, the transmitted data is divided into

consecutive frames. In each frame, the data stream is further divided into $h \times w$ data blocks and modulated by COBRA color code, where h is the column number of the data blocks and w is the row number of the data blocks. Each block occupies $bs \times bs$ pixels, where bs represents the block size (BS) and is mapped to certain color in the color code modulation method. Figure 1.6 shows an example frame in a COBRA system encoded in four colors (white, red, green and blue). Additionally, four corner trackers, which are blocks in black surrounded by eight blocks in the same color of red, green or blue, are added on the corners of each frame to help determine the orientation of the captured image at the receiver side. Furthermore, four timing reference blocks, which are black blocks surrounded by eight white blocks, are added on the margin of the image and used to determine the location of the color data blocks. Then all the $h \times w$ color data blocks, four corner trackers and four timing reference blocks will form one image and display at the transmitter screen as shown in Figure 1.6.

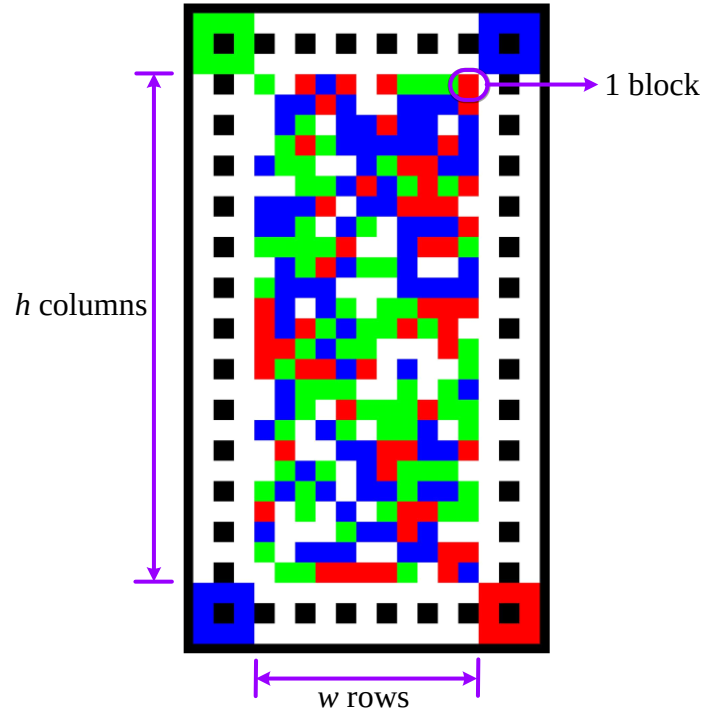


Figure 1.6: An example for a 2D COBRA barcode.

Unfortunately, due to the broadcast property and visual nature of display screens, SCC systems are subject to eavesdropping and malicious attack, especially in public venues, such as supermarkets and shopping malls. This vulnerability has led to a number of recent studies into the security of SCC systems. A physical security enhancement method for barcode-based

SCC systems was proposed in [11]. It can realize secure communication by manipulating screen view angles and leveraging user-induced motions. Kaleido precodes the distortion in the video to prevent unauthorized video recording while maintaining the viewing quality of human eyes [12]. In [13], three secure communication schemes for SCC systems were proposed and all three methods require strict synchronization. A color shift-based secret key distribution scheme was proposed in [14]. A covert SCC system was proposed to embed user data into a video and realize simultaneous display and communication on the screen in [15]. The system can only achieve a 10^{-1} -level bit error rate (BER) when the block size is up to 60 px. An upgraded color barcode scheme for secure transmission in SCC systems was proposed in [16] to improve the throughput based on an advanced smartphone screen and a camera with high resolution. A screen camera-based secure short-range communication system was proposed in [17]. It achieves secure communication links with the aid of the color shift and perspective distortion characteristics of the SCC channel. Leveraging the color shift property over screen-to-camera channel, two key distribution methods was proposed in [18].

Furthermore, location-based services have seen incredible expansion in recent years. Indoor VLC systems can also be used for precise positioning, which is also called visible light positioning (VLP) and will be detailed in CHAPTER 2.

1.3.2 Outdoor VLC Systems

Existing researches on outdoor VLC systems mainly focus on Intelligent Transportation Systems (ITS), including vehicle to vehicle (V2V) [19] and vehicle to infrastructure (V2I) [20] systems. By enabling VLC technology on mobile devices and integrate it with smart street lighting modules, the high-precision positioning and high-speed data rate performances of VLC technology and the wide distribution of street lightings can intensely contribute to the implement of diverse public services. Furthermore, in-vehicle network applications can leverage VLC-enabled lighting modules and traffic infrastructure to implement Vehicular VLC [21].

However, outdoor applications face more challenges due to the strong ambient light disturbance [22] and optical defects in lens. The power of the incident parasitic light can be up

to 10 mW/cm^2 , compared to the power of the light containing the information which can be as low as few $\mu\text{W/cm}^2$. Meteorological phenomena such as rain, snow, fog, and other particles in the atmosphere scatter and absorb light beams, thereby reducing the quality of VLC signals [23]. Furthermore, the accumulated dirt or ice over the transceivers of outdoor VLC systems will obstruct the transmitted signals. These blockages severely affects the reliability, robustness and transmission distance of outdoor VLC. In addition, ice on the road and pavement leads to stronger reflections. Addressing and solving these challenges enables VLC technology in outdoor application and leads to the perspective of fully exploiting the advantages of VLC [22].

1.4 Related Works

Recently, VLC has attracted increasing attention from academia and industry and the number of research works, including exhaustive surveys, on this topic has grown exponentially [3]. The very first work that utilized LEDs to provide illumination and communication services in indoor scenarios as proposed by the scholars from Keio University in Japan in 2000 [24]. In 2003, the Visible Light Communication Consortium (VLCC) was founded in Japan, and VLC technology for mobile devices and vehicles has been rapidly developed.

In addition, standardization work has been carried out by the Japan Electronics and Information Technology Industries Association (JEITA) and the IEEE Standards Association (IEEE-SA). In 2007, VLCC introduced two standards given by the visible light communication system standard and the visible light ID system standard. Then JEITA accepted these two standards in the forms of JEITA CP-1221 [26] and CP-1222 [27], respectively. In 2011, IEEE-SA developed the first IEEE standard for VLC given by IEEE 802.15.7 [25] to define the physical layer (PHY) and medium access control (MAC) layer design specifications.

At the same time, VLC-enabled infrastructures from companies, such as pureLiFi, Philips, Oledcomm, have appeared on the market and are deployed in domestic and industrial buildings [28].

1.5 Thesis Organization

The rest of the thesis is organized as follows. The design of a VLC transmitter system is discussed in Chapter 2. A universal VLC modulator design is presented that is integrated with Bluetooth-based wireless connectivity and supports various LED lighting and a wide range of input power for LED drivers. Chapter 3 describes the use of a smartphone camera as a receiver and characterizes the performance of the communication link with respect to various types of lights, colors and smartphone models. Chapter 4 is focused on smart lighting and display applications of VLC. A smart LED lighting installation and control system is presented, followed by the implementation of OCC in LCD displays. Chapter 5 describes the implementation of a high-accuracy indoor positioning system using the transceiver system built in previous chapters. Finally, Chapter 6 concludes the thesis and gives an overview of future work.

1.6 References

- [1] I. Tavakkolnia, D. Cheadle, R. Bian, T. H. Loh and H. Haas, “High speed millimeter-wave and visible light communication with off-the-shelf components,” in *Proc. 2020 IEEE Globecom Workshops (GC Wkshps)*, Taipei, Taiwan, 2020, pp. 1-6.
- [2] H. Haas, J. Elmirghani and I. White, “Optical wireless communication”, *Phil. Trans. R. Soc. A*, vol. 378, no. 20200051, pp. 1-11, Mar 2020.
- [3] Jovicic, J. Li, and T. Richardson, “Visible light communication: Opportunities, challenges and the path to market,” *IEEE Commun. Mag.*, vol. 51, no. 12, pp. 26–32, Dec. 2013.
- [4] J. A. J. Roufs and F. J. J. Blommaert, “Temporal Impulse and Step Responses of the Human Eye Obtained Psychophysically by Means of a Drift-Correction Perturbation Technique,” *Vision Res.*, vol. 21, no. 8, pp. 1203-21, 1981.
- [5] W. A. Cahyadi, Y. H. Chung, Z. Ghassemlooy, and N. B. Hassan, “Optical camera communications: principles, modulations, potential and challenges,” *Electron.*, vol. 9, no. 9, pp. 1339, 2020.

- [6] D. Milovančev, N. Vokić, H. Hübel and B. Schrenk, “Gb/s Visible Light Communication With Low-Cost Receiver Based on Single-Color LED,” *J. Light. Technol.*, vol. 38, no. 12, pp. 3305-3314, 15 June 2020.
- [7] C. Chow, C. Chen and S. Chen, “Enhancement of Signal Performance in LED Visible Light Communications Using Mobile Phone Camera,” *IEEE Photonics J.*, vol. 7, no. 5, pp. 1-7, Oct. 2015, Art no. 7903607.
- [8] P. H. Pathak, X. Feng, P. Hu and P. Mohapatra, “Visible light communication networking and sensing: A survey potential and challenges,” *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2047-2077, 4th Quart. 2015.
- [9] H. Haas, L. Yin, Y. Wang and C. Chen, “What is LiFi?,” *J. Lightw. Technol.*, vol. 34, no. 6, pp. 1533-1544, Mar. 2016.
- [10] T. Hao, R. Zhou, and G. Xing, “COBRA: Color barcode streaming for smartphone systems,” in *Proc. the 10th international conference on Mobile systems, applications, and services*, UK, 2012, pp. 85–98.
- [11] B. Zhang, K. Ren, G. Xing, X. Fu, and C. Wang, “SBVLC: Secure barcode-based visible light communication for smartphones,” in *Proc. IEEE INFOCOM 2014 - IEEE Conference on Computer Communications*, Toronto, Canada, 2014, pp. 2661–2669.
- [12] L. Zhang, C. Bo, J. Hou, X.-Y. Li, Y. Wang, K. Liu, and Y. Liu, “Kaleido: You can watch it but cannot record it,” in *Proc. the 21st Annual International Conference on Mobile Computing and Networking*, 2015, pp. 372–385.
- [13] 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, 2015.
- [14] H. Liu, B. Liu, C. Shi, and Y. Chen, “Secret key distribution leveraging color shift over visible light channel,” in *Proc. 2017 IEEE Conference on Communications and Network Security (CNS)*, Las Vegas, NV, USA: IEEE, 2017, pp. 1–9.

- [15]J. Wang, W. Huang, and Z. Xu, “Demonstration of a covert camera-screen communication system,” in *Proc. 2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC)*, Valencia, Spain: IEEE, 2017, pp. 910–915.
- [16]M. Stafford, A. Rogers, S. Wu, C. Carver, N. S. Artan, and Z. Dong, “TETRIS: Smartphone-to-smartphone screen-based visible light communication,” in *Proc. 2017 IEEE 14th International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*, IEEE, 2017, pp. 570–574.
- [17]J. Zhao and X.-Y. Li, “SCsec: A secure near field communication system via screen camera communication,” *IEEE Trans. Mob. Comput.*, vol. 19, no. 8, pp. 1943–1955, 2019.
- [18]H. Liu, C. Shi, and Y. Chen, “Enabling secret key distribution over screen-to-camera channel leveraging color shift property,” *IEEE Trans. Mob. Comput.*, 2021.
- [19]B. Soner and S. C. Ergen, “Vehicular Visible Light Positioning with a Single Receiver,” in *Proc. 2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2019, pp. 1-6.
- [20]V. K. G. Kalaiselvi, A. Sangavi and Dhivya, “Li-Fi technology in traffic light,” in *Proc. 2017 2nd International Conference on Computing and Communications Technologies (ICCCT)*, 2017, pp. 404-407.
- [21]M. Uysal, Z. Ghassemlooy, A. Bakkali, A. Kadri and H. Menouar, “Visible light communication for vehicular networking: Performance study of a V2V system using a measured headlamp beam pattern model,” *IEEE Veh. Technol. Mag.*, vol. 10, no. 4, pp. 45-53, Dec. 2015.
- [22]Y. Zhuang et al., “A Survey of Positioning Systems Using Visible LED Lights,” *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 1963-1988, 3rd Quart. 2018.
- [23]R. Martinek, L. Danys and R. Jaros, “Visible light communication system based on software defined radio: Performance study of intelligent transportation and indoor applications,” *Electron.*, vol. 8, no. 4, pp. 433, 2019.

- [24]Y. Tanaka, S. Haruyama and M. Nakagawa, “Wireless optical transmissions with white colored LED for wireless home links,” in *Proc. 11th IEEE Int. Symp. PIMRC*, vol. 2, pp. 1325-1329, 2000.
- [25]IEEE, 802.15.7-2011, “IEEE Standard for Local and Metropolitan Area Networks—Part 15.7: Short-Range Wireless Optical Communication Using Visible Light”, Sep. 2011.
- [26]JEITA, CP-1221, “Visible Light Communications System”, Mar. 2007.
- [27]JEITA, CP-1222, “Visible Light ID System”, Jun. 2007.
- [28]A. Memedi and F. Dressler, “Vehicular Visible Light Communications: A Survey,” *IEEE Commun. Surveys Tuts.*, vol. 23, no. 1, pp. 161-181, 1st Quart. 2021.

CHAPTER 2 Visible Light Positioning Systems

2.1 Positioning Technologies

In outdoor scenario, global positioning system (GPS) is widely used to provide location-based services. GPS uses satellite as transmitter to acquire the position of airplanes, automobiles and portable devices. A GPS receiver locate four or more satellites and calculate the distance to each satellite to infer its own position. GPS calculates the distance by measuring satellite electronic clock, and determine the location of receivers through triangulation method [1], based on which only three satellites are needed to estimate the position of a GPS receiver. However, the result is not accurate due to the error of the clock. The fourth satellite is used to estimate the related position to the three satellites to reduce the positioning error. Several studies focused on improving the positioning accuracy of GPS system, such as differential GPS (DGPS) [2] and GPS roadside integrated precision positioning system (GPSIPPS) [3]. However, these systems require high-cost receivers.

In more complex indoor environments, GPS can not achieve high-robustness positioning performance due to multipath propagation, signal blockage and attenuation [4]. Indoor positioning has become a difficult task, and there is no universal solution for all operations [5]. Radio frequency (RF) sensors can help enhance the positioning performance in indoor scenes. Under the consideration of the cost of large-scale employment, Wi-Fi, Bluetooth, radio frequency identification (RFID) are the most common used indoor positioning technologies. Wi-Fi-based indoor positioning system (IPS) was first proposed by Microsoft Research using received signal strength (RSS) method [6], which measures the signal strength between the receiver and multiple access points to estimate the distance using propagation model. Afterwards, most Wi-Fi-based IPS uses fingerprint method. The fingerprint method is also based on RSS, but it stores this knowledge in the database and compares it with the known positions of the receivers. However, it is challenging to achieve high-accuracy positioning results exclusively due to multipath effect [7].

iBeacon is a protocol proposed by Apple based on Bluetooth Low Energy (BLE) devices and aimed to push notifications to nearby visitors via mobile applications when they are approaching the BLE devices [8]. BLE devices broadcast their identifier (ID) to nearby mobile devices to achieve position acquiring and tracking. Compared with traditional Bluetooth, BLE requires lower power consumption, lower cost and less amount of data to be transmitted periodically.

As a cost-effective item serialization wireless technology, RFID is widely applied in retail scenarios and RFID-based sensing network focuses on providing location-based cloud services [9]. RFID reader and RFID tag are two basic components in RFID systems. RFID tags contain antenna-based transceivers and integrated circuit (IC) for RF signal modulation [10]. RFID tags are always attached to objects to be located by RFID readers. RFID readers transmit power to RFID tags, so there is no batteries in a RFID tag. As the signal strength of RFID system is essentially influenced by noise, multipath effect, and antenna interference, the arrangement of RFID tags is still a challenging task. Support vector machine (SVM) technology can be applied to RFID-based IPS to improve the positioning precision [11]. As described above, different indoor positioning technologies have different features and different applications. The summary of different RF-based IPSs [12] is given in Table 2.1.

Table 2.1: Summary of RF-based IPS [12].

Technology	Coverage	Accuracy	Scalability	Cost	Power Consumption
GPS	16 km	6-20 m	Low	High	High
Wi-Fi	35 m	1-5 m	Medium	Medium	High
Bluetooth	10 m	1-5 m	High	Low	Low
RFID	1 m	1-2 m	Medium	Low	Low

Non-RF technologies can also realize localization, such as magnetic positioning, pedestrian dead reckoning (PDR) and image-based positioning. A magnetic positioning system usually measures magnetic field strength from a permanent magnet to a magnetic sensor, which is usually employed for intelligent industrial applications, such as automobile shift detection, wheel speed sensing [13]. It has multiple advantages including high reliability, high positioning

accuracy and low manufacturing cost. A practical magnetic system design must analyze the influence of stray field, ferromagnetic environment, limited installation space, required resolution, compensation of manufacturing tolerance, leading to high mathematical challenges [14].

PDR technology calculates the number of steps based on the built-in sensors of mobile devices to estimate people's positions [15]. It can only achieve coarse positioning accuracy because it is challenging to precisely estimate the exact step size of the pedestrian and heading direction. The walking mode and step size of the pedestrians are varying and the heading direction measured by compass is sensitive to around electrical devices. Therefore, other IPS technologies are applied to PDR-based positioning systems to estimate the accumulated errors, such as Wi-Fi, Bluetooth, magnetism or acoustics.

Among the above technologies, VLP technology is very competitive in providing indoor localization services due to its high availability, high bandwidth, low cost and long lifetime. VLP signals cannot penetrate through walls, ceilings and other obstructions inside buildings. Thus, different VLP systems will not cause interference to each other and the signals in different rooms are independent and private. VLP technology provide positioning with centimeter-level accuracy and illumination for large smart manufactories, high-rise buildings, densely populated shopping malls and underground environments. It is noteworthy that VLP systems are scalable especially for multi-floor buildings and can make the two positioning systems for robots and mobile devices, respectively, share the same map and achieve positioning on the same map.

2.2 VLP System Configuration

2.2.1 PD-based VLP Systems

The most conventional transceiver in PD-based VLP system consists of one PD as receiver and multiple LEDs as transmitters. An example of a single PD-based VLP system with four LEDs as transmitters is illustrated in Figure 2.1. As shown in Figure 2.1, a PD-based receiver is placed on the receiver plane and four LEDs are mounted on the ceiling. Four LEDs transmit

different signals given by f_1 , f_2 , f_3 and f_4 . Then at the receiver side, the RSS, the time of arrival (TOA) or the angle of arrival (AOA) will be measured.

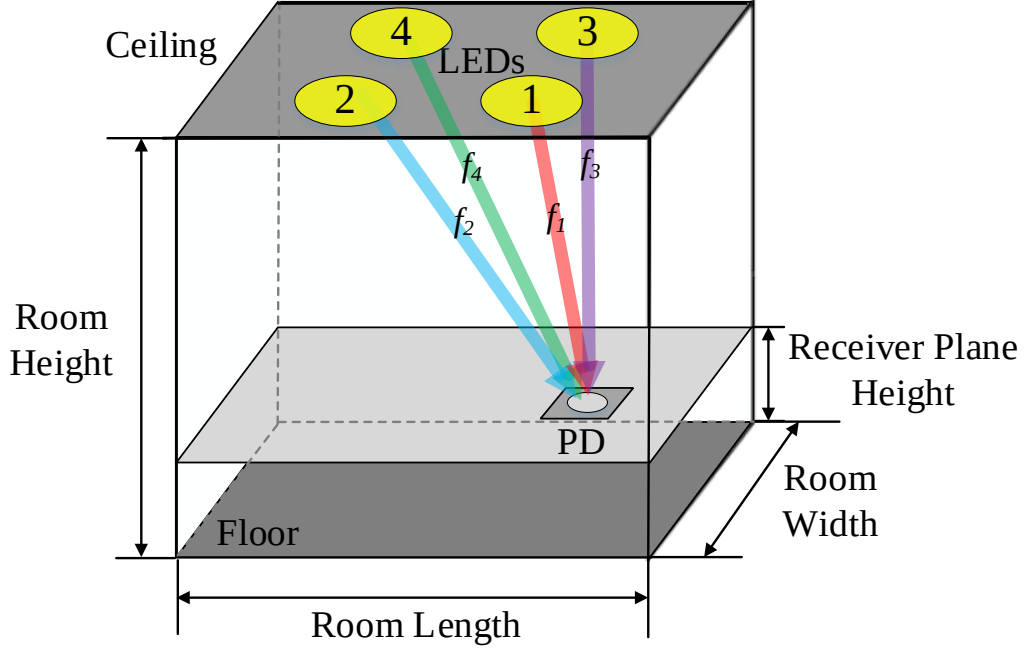


Figure 2.1: An indoor VLP system based on a single PD and four LEDs.

RSS-based positioning system is easily employed without the need for any complementary equipment except for a single PD and multiple LEDs. Therefore, it is more cost-effective than TOA-based VLP systems which require additional devices to estimate time. In a RSS-based VLP system, the target location can be estimated by the received direct current (DC) signal strength or the output alternating current (AC) signal strength of the PD. However, both of these methods have potential shortcomings that received DC signal strength-based systems are sensitive to DC measurement error caused by ambient light and output AC signal strength-based systems require perfect transceiver parameters [16]. Many conventional algorithms can be applied to RSS-based VLP system, including trilateral evaluation [16], fingerprint perception [17] and proximity [18].

In addition, a time difference of arrival (TDOA)-based VLP system measures the traveling time of the light wave from the LED-based transmitters to the PD-based receiver. It requires at least three LEDs to achieve 2D positioning and four LEDs to realize 3D positioning. However, precise time synchronization between the LEDs and the PDs is a prerequisite, which will

increase the manufacturing cost of the system. With the help of multiple LEDs modulated with frequency division multiplexing (FDM) and band pass filter (BPF) applied on the receiver side [19], a single PD can collect optical signals independently and measure the time difference between the signals, which facilitates the practical implementation of TDOA-based VLP system.

An AOA system measures the direction of propagation of the light signal incident on the PD. In an AOA-based VLP system, the PD-based receiver is surrounded by multiple sector LEDs with limited field of view (FOV) [20]. The position of the PD can be approximately estimated using AOA method, when at least two optical signals from different transmitters are received. By mounting more LEDs in the AOA-based VLP system, higher positioning accuracy can be achieved. Table 2.2 summaries the features of the PD-based VLP systems using different positioning methods [12].

Table 2.2: Comparison of different PD-based VLP systems [12].

Positioning Methods	Accuracy	Cost	Power Consumption	Multipath Effect
RSS	Medium	Medium	Low	Yes
TDOA	High	High	Low	Yes
AOA	Medium	Low	High	Yes

2.2.2 Image Sensor-based VLP Systems

A smart phone based-VLP system consists of an LED based transmitter and the image sensor as receiver. The front-end design of the VLC encoding and decoding is illustrated in Figure 2.2. The LED is controlled by a VLP enabled LED driver to transmit optical waveforms. Then at the receiver side, the CMOS image sensor on the smart phone works at rolling shutter mode and the pixel rows are to be captured from top to bottom [23]. Thus, as long as the data rate of the LED is modulated higher than the frame rate of the smart phone, the transmitter signals will be recorded as the rolling shutter patterns on the image sensor. To present straightforwardly, we assume that the transmitter applies on-off keying (OOK) modulation. Bright and dark bars correspond to transmitted data 1 and 0 as illustrated in Figure 2.2(b). The LED is first turned off, which results in a lower intensity output on the first-column pixels at the receiver camera

side. Then the LED driver control the LED by switching it to the on state, which will be recorded as the higher intensity output on the second-column pixels. Therefore, expert guidance is required when setting up a VLP system. In the proposed VLP system, the receiver camera works at rolling shutter mode to achieve a higher data rate by reading the PD matrix in the camera column by column compared with that working at global shutter mode. After scanning all columns, the image sensor-based receiver will convert all the columns on the resultant image into binary data. By working at the rolling shutter mode, the image sensor-based VLP system can achieve multi-kbps throughput. Therefore, rolling shutter processing can be used to increase the data rate. The image sensor-based receiver decodes the patterns based on threshold. After adding a preamble and error check sequence, the decoded binary sequences will consist a data frame with unique ID and then the ID is mapped to a uniform resource identifier (URI) database and read by the application software [22].

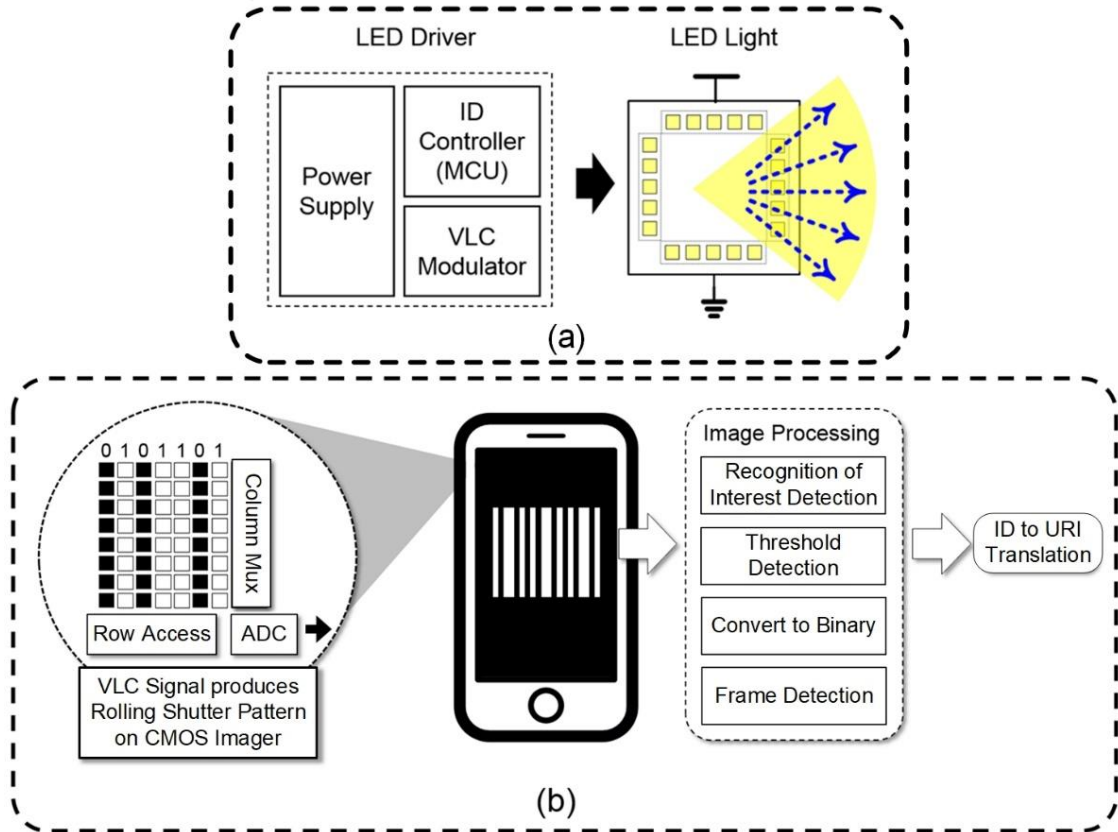


Figure 2.2: Block diagram of a smart phone based VLC system illustrating (a) transmitter side and (b) receiver side [22].

To acquire the position of an object, an image sensor-based VLP system usually relies on an auxiliary magnetometer to estimate the azimuth angle. In [21], a double image sensors-based VLP system without angle estimation is proposed, where two equivalent image sensors are placed horizontally with the same height and trilateral evaluation is applied. However, double image-sensors-based VLP systems require high cost and additional arrangement. Therefore, VLP systems based on a single image sensor is widely studied.

2.3 VLP Technology Applications

2.3.1 Location-aware Services

It is easy to get lost for humans in indoor public environments, such as shopping malls, airports, museums and exhibition centers. Location-aware services will be a precious asset to lead the way for us to our areas of interest [24]. Since LED lightings are already mounted in these public areas for illumination, VLP systems can be easily facilitated based on the existing lighting infrastructures. It is noteworthy that certain adjustment must be added to the original equipment manufacturers (OEMs) of the LEDs to generate VLC signals. The VLP-enabled smartphones or other mobile devices can demonstrate the information of the around exhibitions, the advertisement of nearby stores and guide people to the nearest toilet, elevator or ATM. In a multi-story building, VLP can even achieve 3D positioning by encoding the 3D locations in the LEDs. In addition, if certain area is getting overcrowded, especially for pandemic control under covid-19, VLP systems can help the staffs to regulate the flow of people.

2.3.2 Robotics Navigation and Localization

In recent years, Industry 4.0 is revolutionizing the way manufactories product goods. To meet the demanding challenge of the Industry 4.0 application requirement, increasing amount of RFID devices, wireless networks and sensors are integrated in the smart manufacturing and logistics distribution centers. Mobile robots are also assembled on production lines to realize manufactory automation. However, the communication reliability is limited due to some potential problems of indoor wireless communication systems, such as multipath propagation, shadow effect, signal attenuation and interference [25]. In addition, most of existing industrial

mobile robots are heavily limited to predetermined and fixed route without the capability of intelligent positioning and self-navigation.

In intelligent manufactories and industrial logistics centers, VLP technology can be applied to navigate mobile robots to improve work efficiency when they are controlled to conduct resource allocation and regulate repository management [26]. Beyond Industry 4.0, Industry 5.0 is projected to extend the capability of multi-robot collaboration and human-machine cooperation and interaction, as shown in Figure 2.3. In addition, by encoding different LEDs with different IDs, VLP systems can cover the entire workshop to provide scalable location-based services and navigate the robots to handle dynamic tasks by rearranging the navigation paths and destinations for robots. Besides industrial environments, mobile service robots are also widely used in public places, such as museums, airports, shopping malls and hospitals to realize information display, path guide, ward patrol and room cleaning.

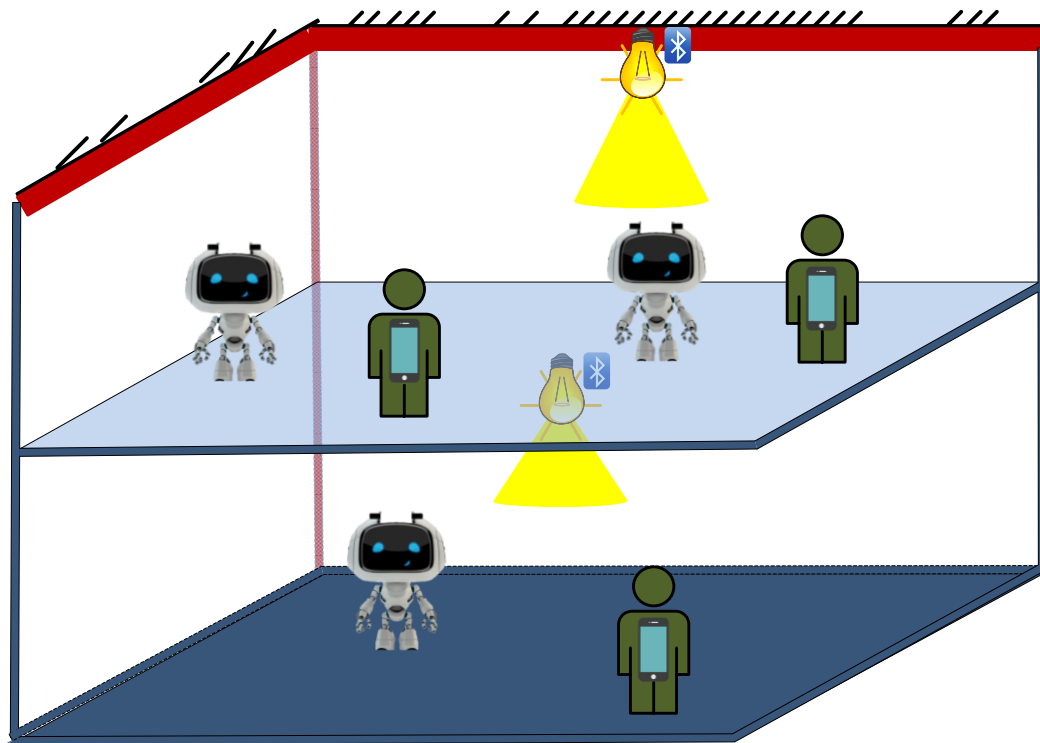


Figure 2.3: Diagram of a common indoor VLP system with multiple mobile robot collaboration and human-robot cooperation under based on ceiling mounted LED lightings.

Nowadays, Simultaneous Localization and Mapping (SLAM) technology has been considered as the key for robots to achieve autonomous positioning and navigation, which was

first proposed by John J. Leonard and Hugh F. Durrant-Whyte in 1992 [27]. SLAM is mainly used to solve the problems of positioning, navigation and map construction when mobile robot runs in an unknown environment and provides various services, such as food delivery, shopping guide and bank self-service. Table 2.3 illustrates several key benefits of VLP technology in mapping, deployment, and robustness when compared with current SLAM-based positioning methods.

Table 2.3: Advantages of VLP in comparison to SLAM-based positioning systems.

Approach	VLP	SLAM
Mapping process	Do not require scanning of venue	Require scanning and survey for positioning
Map share	Just the building floorplan in BIM with lights' location is enough	Challenging to share the same map between different robots
Equipment	Modulated LED	Sophisticated and high-cost sensors
Scalability	Global 3D	Local 2D/3D
Environment dependency	Not based on features	Poor performance in repeatable feature environment and featureless area

2.4 References

- [1] McNeff, J.G, “The global positioning system,” *IEEE Trans. Microw. Theory Tech.*, vol.50, pp. 645 - 652, March 2002.
- [2] R. S. Rempel and A. R. Rodgers, “Effects of differential correction on accuracy of a GPS animal location system”, *J. Wildl. Manage.*, vol. 61, no. 2, pp. 525-530, Apr. 1997.
- [3] D. Hohman and T. Murdock, E. Westerfield, T. Hattox and T. Kusterer, “GPS roadside integrated precision positioning system”, in *Proc. IEEE 2000. Position Location and Navigation Symposium (Cat. No.00CH37062)*, pp. 221-230, 2000.
- [4] M. M. Atia, A. Noureldin and J. M. Korenberg, “Dynamic online-calibrated radio maps for indoor positioning in wireless local area networks”, *IEEE Trans. Mobile Comput.*, vol. 12, no. 9, pp. 1774-1787, Sep. 2013.

- [5] R. Joseph and S. B. Sasi, "Indoor Positioning Using WiFi Fingerprint," in *Proc. 2018 International Conference on Circuits and Systems in Digital Enterprise Technology (ICCSDET)*, 2018, pp. 1-3.
- [6] P. Bahl and V. N. Padmanabhan, "RADAR: an in-building RF-based user location and tracking system," in *Proc. IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No.00CH37064)*, 2000, pp. 775-784, vol.2.
- [7] Z. Wang, P. Sokliep, C. Xu, J. Huang, L. Lu and Z. Shi, "Indoor Position Algorithm Based on the Fusion of Wifi and Image," in *Proc. 2019 Eleventh International Conference on Advanced Computational Intelligence (ICACI)*, 2019, pp. 212-216.
- [8] L. Yu, Y. Liu, T. Chi and L. Peng, "An iBeacon-based indoor and outdoor positioning system for the fire emergency command," in *Proc. 2017 Forum on Cooperative Positioning and Service (CPGPS)*, 2017, pp. 326-329.
- [9] H. Mora-Mora, V. Gilart-Iglesias, D. Gil and A. Sirvent-Llamas, "A computational architecture based on RFID sensors for traceability in smart cities", *Sens.*, vol. 15, pp. 13591-13626, 2015.
- [10] M. Sasikala, J. Athena and A. S. Rini, "Received Signal Strength based Indoor Positioning with RFID," in *Proc. 2021 IEEE International Conference on RFID Technology and Applications (RFID-TA)*, 2021, pp. 260-263.
- [11] He Xu et al., "An RFID indoor positioning algorithm based on support vector regression", *Sens.*, vol. 18.5, pp. 1504, 2018.
- [12] J. Luo, L. Fan and H. Li, "Indoor Positioning Systems Based on Visible Light Communication: State of the Art," *IEEE Commun. Surv. Tutor.*, vol. 19, no. 4, pp. 2871-2893, Fourthquarter 2017
- [13] M. Ortner, M. Ribeiro and D. Spitzer, "Absolute Long-Range Linear Position System With a Single 3-D Magnetic Field Sensor," *IEEE Trans. Magn.*, vol. 55, no. 1, pp. 1-4, Jan. 2019, Art no. 4000104.

- [14]C. Huber et al., “Topology optimized and 3d printed polymer-bonded permanent magnets for a predefined external field,” *J. Appl. Phys.*, vol. 122, no. 5, pp. 053904, 2017.
- [15]Q. Tian, Z. Salcic, K. Wang, and Y. Pan, “A multi-mode dead reckoning system for pedestrian tracking using smartphones,” *IEEE Sensors J.*, vol. 16, no. 7, pp. 2079–2093, Apr. 2016.
- [16]N. Huang, C. Gong, J. Luo and Z. Xu, "Design and Demonstration of Robust Visible Light Positioning Based on Received Signal Strength," *J. Lightw. Technol.*, vol. 38, no. 20, pp. 5695-5707, 15 Oct.15, 2020.
- [17]U. Nadeem, N. U. Hassan, M. A. Pasha and C. Yuen, “Indoor positioning system designs using visible LED lights: Performance comparison of TDM and FDM protocols,” *Electron. Lett.*, vol. 51, no. 1, pp. 72-74, Jan. 2015.
- [18]Y. Zhao et al., “Proximity report triggering threshold optimization for network-based indoor positioning,” in *Proc. 2015 18th International Conference on Information Fusion (Fusion)*, pp. 1061-1069, 2015.
- [19]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, Nov. 2011.
- [20]Y. S. Erogluy, I. Guvency, N. Palay and M. Yukselz, “AOA-based localization and tracking in multi-element VLC systems,” in *Proc. 2015 IEEE 16th Annual Wireless and Microwave Technology Conference (WAMICON)*, pp. 1-5, 2015.
- [21]M.-G. Moon, S.-I. Choi, J. Park and J. Y. Kim, “Indoor positioning system using LED lights and a dual image sensor,” *J. Opt. Soc. Korea*, vol. 19, no. 6, pp. 586-591, Dec. 2015.
- [22]B. Hussain, C. Lau and C. P. Yue, “Li-Fi based Secure Programmable QR Code (LiQR),” *JSAP-OSA Joint Symposia*, 2017, pp. 6p_A409_6.
- [23]C. W. Chow, C. Y. Chen, and S. H. Chen, “Enhancement of signal performance in LED visible light communications using mobile phone camera,” *IEEE Photon. J.*, vol. 7, no. 5, Oct. 2015, Art. no. 7903607.

- [24]Y. Zhuang, Z. Syed, J. Georgy and N. El-Sheimy, “Autonomous smartphone-based WiFi positioning system by using access points localization and crowdsourcing,” *Pervasive Mobile Comput.*, vol. 18, pp. 118-136, Apr. 2015.
- [25]Willig, “Recent and emerging topics in wireless industrial communications: A selection,” *IEEE Trans. Ind. Informat.*, vol. 4, no. 2, pp. 102-124, May 2008.
- [26]S. Chang and C. Lin, “A Visible Light Communication Positioning Mechanism in Industrial Logistics Management,” in *Proc. 2016 30th International Conference on Advanced Information Networking and Applications Workshops (WAINA)*, 2016, pp. 878-882.
- [27]John J. Leonard, Hugh F. Durrant-Whyte and Ingemar J Cox, “Dynamic map building for an autonomous mobile robot,” *Int. J. Rob. Res.*, vol. 11, no.4, pp. 286-298, 1992.

CHAPTER 3 High-Accuracy Indoor VLP Systems for Mobile Devices

3.1 Introduction

With the growing demand for location-based services such as indoor navigation, robot control and object tracking, indoor positioning technology has attracted increasing attention from both academia and industry. For outdoor environments, the Global Positioning System (GPS) provides real-time positioning services based on satellites and is widely used in airplanes, automobiles and portable devices. However, it cannot realize efficient positioning in indoor environments because satellite signals will be extremely attenuated and interrupted by indoor obstacles [1]. Currently, wireless technologies, including Bluetooth [2] and WiFi [3], are widely applied to indoor positioning systems. However, these technologies can only achieve meter-level accuracy [4] and are potentially vulnerable to malicious activities [5]. Visible light positioning (VLP) technology can solve these problems, with multiple advantages including centimeter-level accuracy, compatibility with existing lighting infrastructure, low cost and insusceptibility to electromagnetic interference. Therefore, VLP systems are very competitive to provide indoor positioning service.

VLP systems use light emitting diodes (LEDs) as transmitters and have two different types of receivers, image sensors and photodiodes (PDs). The detection areas of PDs are very small. Thus, precise alignment between the transmitter and a PD-based receiver is required for signal detection [6]. Compared with PDs, image sensors are more widely incorporated into mobile devices, promoting their universal deployment. Additionally, a lot of work has been done on smartphone camera-based visible light communication (VLC) [7] and robot camera-based VLC [8]. In an image sensor-based VLP system, the LED-based transmitter is controlled by a VLC-enabled LED driver to transmit optical waveforms and the CMOS image sensor works in rolling shutter mode [9]. The receiver decodes the patterns based on a threshold, and the decoded binary sequences will consist of a data frame with unique identification (ID) and be mapped to a uniform resource identifier (URI) database [10].

In this chapter, we propose a tilted receiver camera correction and partially blocked LED image compensation algorithm to realize high-accuracy positioning for VLP systems. The proposed positioning method is based on a single LED, and the additional positioning error caused by tilted receiver is corrected by the rotation angles estimated by the sensors on the smartphones. Additionally, the LED-based transmitters are modulated with digital IDs containing the information of the LEDs' world coordinates. Therefore, the positioning method is scalable, with no maximum scale limit. For example, in a multi-floor multi-room building, by encoding the LEDs in different rooms on different floors with different IDs, VLP systems can provide 3D positioning in the venue. The proposed VLP system can also provide positioning services even when an incomplete LED image is captured by the camera. Two different methods are proposed to find the geometric features when only part of the LED curve is shown on the captured images. With these methods, the robustness of the VLP system is enhanced. We finally present experiments to verify the effectiveness of the proposed VLP methods and analyze the performance.

The chapter is organized as follows. Related work is introduced in Section 3.2. The proposed tilted receiver camera correction method for VLP systems is introduced in Section 3.3. In Section 3.4, we present the details of the proposed partially blocked LED image compensation method. Experimental results are provided and analyzed in Section 3.5. Finally, Section 3.6 concludes this chapter.

3.2 Related Works

During recent years, a significant amount of work has been done on VLP systems using image sensors as receivers. A novel VLP system based on an event-based neuromorphic vision sensor was proposed in [11]. It achieves positioning error lower than 3 cm when the height between the LEDs and the event camera is within 1 m. However, it is noteworthy that most of the existing image sensor-based VLP systems have an assumed prerequisite that the image sensor must be placed horizontally, i.e., parallel with the ground. However, in realistic scenarios, smartphones held in the human hand arbitrarily rotate about the x-, y- and z-axes, and rotation correction methods have to be applied to suppress additional positioning errors induced by the

tilt angles. A novel sensor fusion method was presented in [12], which proposes to collect the tilt angle data from the sensors (accelerometer and gyroscope) on a smartphone and suppress the errors caused by the rotation, and the average positioning error is 4.3 cm. An image sensor noise degradation mechanism was proposed in [13], and it uses an accelerometer to measure the tilt angles. The average positioning error is 10 cm when the distance from the LED and the image sensor is 3.5 m. A machine learning method for tilt angle correction was proposed in [14]. Because different tilt angles lead to different characteristics of the LED image, it used neural networks to establish the relationship between the LED image characteristics and the distance between the receiver and the transmitter and correct the error induced by tilted angles, and the average positioning error can reach 1.9 cm. However, the VLP methods proposed in [12-14] use multiple LEDs, which requires the LEDs to be placed at high density. Therefore, to add less tense requirements to the venue to be perceived, single LED-based VLP system is considered in this thesis. A single-LED VLP system with a marker on the LED was proposed in [15], which uses the geometric features of the captured LED images to correct the tilt angles of the camera. However, the x-y plane average 2D positioning error can reach 17.52 cm. A sensor-aided single-LED VLP system was proposed in [16], which relies on the geomagnetic field sensor and accelerometer on a smartphone to estimate the rotation angles, and thereby reduce the positioning error caused by the rotation. However, it only provided the experimental results when the camera is placed directly below the LED and the mean positioning error is about 10 cm when the distance between the LED and the camera is 2.4 m. An orientation calibration method for a single LED-based VLP system was proposed by [17]. However, its average positioning error reaches up to 11.2 cm at the experimental scale of $1.8 \times 1.8 \times 1 \text{ m}^3$. A geometric feature-based planes intersection-line scheme was proposed in [18] and it can achieve the average positioning error of 5.58 cm in the experimental scale of $2.7 \times 1.8 \times 1.45 \text{ m}^3$. However, all of the above VLP methods can achieve high precision when a complete LED image is captured.

A further challenge of VLP systems in practical scenarios is that the cameras are not always able to capture complete LED images. For example, when a camera is placed on the margin of

the field of view (FOV), it can only capture an incomplete LED image. Additionally, when someone is standing below the LED and holding a camera to seek his location, his head partially blocks the LED light. Similarly, when a robot is controlled to track an object in an intelligent manufactory, it may move under tables or shelves and the furniture will partly obstruct the LED light. The resulting incomplete captured LED images will lead to additional positioning error and degrade the robustness of VLP systems. Furthermore, due to a limited FOV, a VLP method based on incomplete LED images can achieve positioning on the margin of the FOV, and can therefore extend the positioning area. By combining a tilted receiver camera, compensation on incomplete LED images can preserve a stable positioning performance on the margin of the extended positioning area. A VLP method based on the mean shift algorithm and unscented Kalman filter was proposed in [19], which enhances the robustness of the VLP system. The accuracy is maintained even when half of the LED is blocked. A Camshift algorithm combined with Kalman filter was proposed in [20] to realize good robustness. However, neither of them considered the positioning error caused by tilted receiver cameras.

3.3 Arbitrarily Tilted Receiver Camera Correction Method for VLP Systems

The diagram of the proposed single circular LED-based positioning method is given in Figure 3.1. In the proposed indoor VLP system, the LEDs are modulated with unique digital IDs containing their own position information, which is stored in a database. The LEDs broadcast their IDs repeatedly. A user holds a smartphone and the front camera captures images of a LED periodically. Then two procedures are executed. The first is to decode the rolling shutter patterns shown on the captured images, find the unique ID of the LED and then obtain the world coordinates of the LED. The other is to detect the boundary of the LED on the captured images and then compute the translation matrix of the camera. The system will also estimate the tilt angles using the sensors on the smartphone and thereby compute the rotation matrix of the camera. Finally, the world coordinates of the camera will be determined.

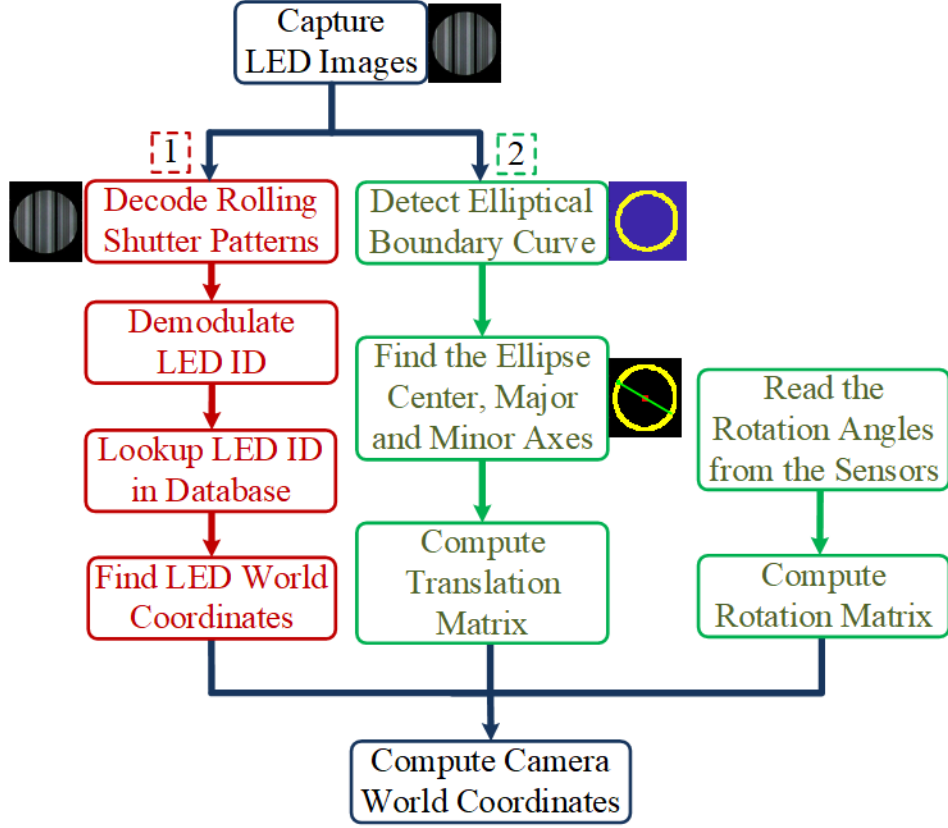


Figure 3.1: The proposed tilted receiver camera correction method for a single-LED-based VLP system: (1) finding the LED’s world coordinates, (2) computing the camera matrices.

3.3.1 Camera Model and Imaging Geometry

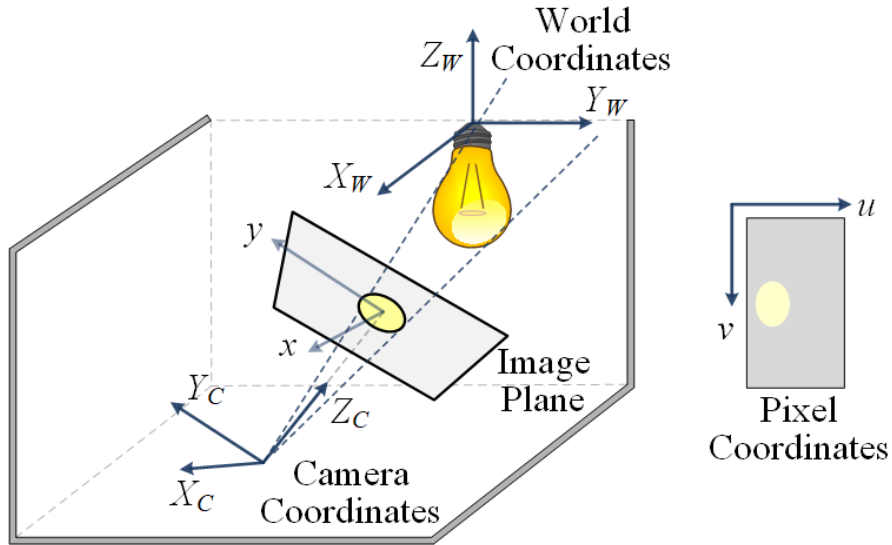


Figure 3.2: Imaging geometry of a single-LED-based VLP system with a tilted camera.

When an image sensor captures a picture of an object, the 3D world coordinates (X_W , Y_W , Z_W) of the object are projected to 2D pixel coordinates (u , v), as shown in Figure 3.2. In the

projection procedure, the world coordinates are firstly converted to camera coordinates (X_C , Y_C , Z_C). Then the 3D camera coordinates are projected to 2D image coordinates (x , y) on the image plane and finally converted to 2D pixel coordinates (u , v). The proposed VLP system realizes navigation by establishing a mapping between the world coordinates and corresponding pixel coordinates.

The camera model describes the mapping from 3D world coordinates to 2D pixel coordinates and is realized by a multiplication of a 3×4 intrinsic matrix $\mathbf{P}_{in}$ and a 4×4 extrinsic matrix $\mathbf{P}_{ex}$. The intrinsic matrix describes the camera's internal parameters, and the extrinsic matrix describes the camera's location and direction in the world coordinate system given by

$$\mathbf{P}_{in} = \begin{bmatrix} f_x & \gamma & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad (3-1)$$

$$\mathbf{P}_{ex} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ \mathbf{0} & 1 \end{bmatrix}, \quad (3-2)$$

where (u_0, v_0) are the pixel coordinates of the principal point, namely, the center of the image, γ is the skew coefficient between the x- and y-axes, often 0, f_x is the focal length of the x-axis in terms of pixels, f_y is the focal length of the y-axis in terms of pixels, $\mathbf{R}$ is the 3×3 rotation matrix and $\mathbf{T}$ is the 3×1 translation matrix. The intrinsic matrix can be obtained via camera calibration [21]. Using these parameters, the projection mapping from world coordinates (X_W , Y_W , Z_W) to pixel coordinates (u , v) can be described as

$$z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & \gamma & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix}, \quad (3-3)$$

where z_c is the scaling factor.

The translation matrix in (3-2) consists of the translation of the camera on the x-, y- and z-axes, respectively, given by

$$\mathbf{T} = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}. \quad (3-4)$$

The rotation matrix in (3-2) is the multiplication of the three rotation matrices describing the rotation of the z-, x- and y-axes, respectively, given by

$$\begin{aligned}
 \mathbf{R} &= \mathbf{R}_z(\varphi_z)\mathbf{R}_x(\varphi_x)\mathbf{R}_y(\varphi_y) \\
 &= \begin{bmatrix} \cos\varphi_z & \sin\varphi_z & 0 \\ \sin\varphi_z & \cos\varphi_z & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi_x & \sin\varphi_x \\ 0 & -\sin\varphi_x & \cos\varphi_x \end{bmatrix} \\
 &\quad \times \begin{bmatrix} \cos\varphi_y & 0 & \sin\varphi_y \\ 0 & 1 & 0 \\ \sin\varphi_y & 0 & \cos\varphi_y \end{bmatrix}, \tag{3-5}
 \end{aligned}$$

where φ_z is the azimuth angle between the y-axis and true north, φ_x is the pitch angle describing the rotation about the x-axis and φ_y is the roll angle describing the rotation about the y-axis, as illustrated in Figure 3.3.

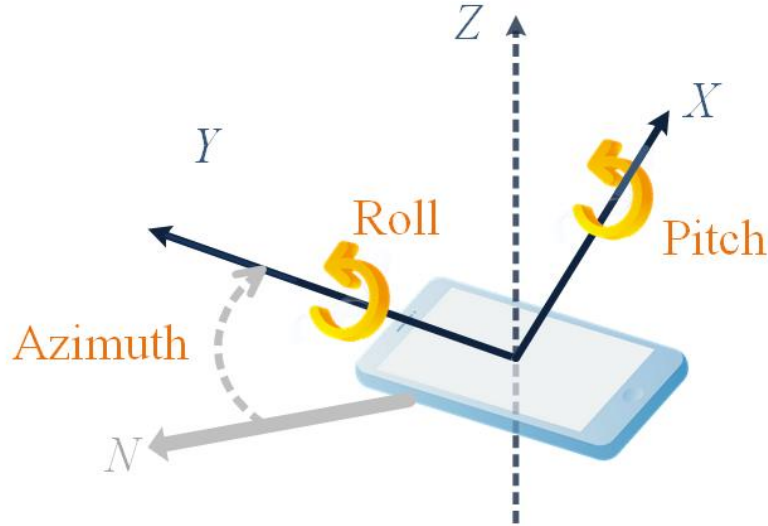


Figure 3.3: Rotation angles of a smartphone along x-, y- and z-axes.

3.3.2 Tilted Receiver Camera Correction Method

When a circular LED is mounted on the ceiling and the image sensor is placed horizontally, that is, the image plane is parallel with the ground, a circular LED image will be captured. However, when the image sensor is rotated about the x-, y- or z-axis, an elliptical LED image will be received. Therefore, LED images captured at the same position may present different geometric features. Fig. 4 shows four images captured at the same position but with different rotation angles. When the tilt angle changes, the pixel coordinates of the LED image and the major and minor axes of the elliptical LED image will also change.

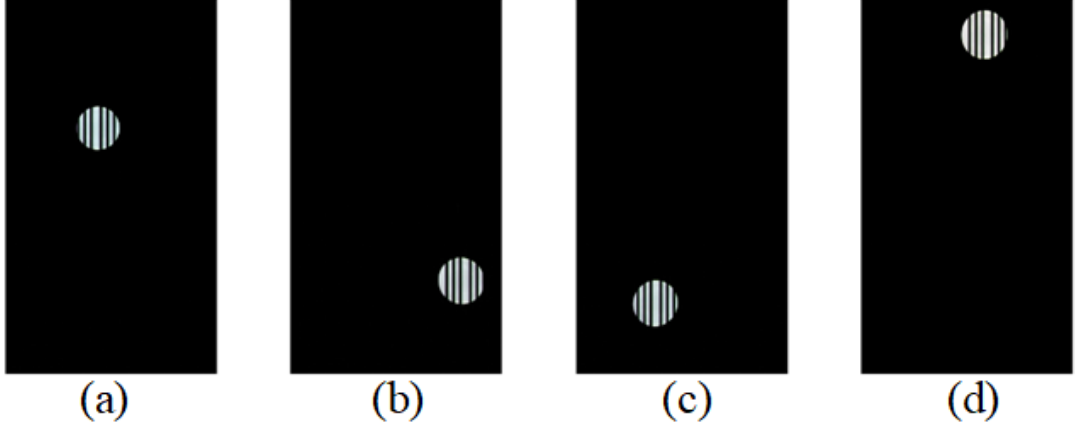


Figure 3.4: Images captured at the same position with different rotation angles: (a) $\varphi_x = 0^\circ$, $\varphi_y = 0^\circ$; (b) $\varphi_x = 15^\circ$, $\varphi_y = -9^\circ$; (c) $\varphi_x = 17^\circ$, $\varphi_y = 3^\circ$; (d) $\varphi_x = -20^\circ$, $\varphi_y = -2^\circ$.

Since the proposed VLP method is based on smartphone images, we take advantage of the accelerometer and geomagnetic field sensor on the smartphone [22], to directly and promptly obtain the tilt angles. The tilt angles are available as the sensor application programming interface (API) through Android and iOS operating systems for mobile devices. Similarly for robots, these readings can be acquired from the inertial measurement unit (IMU). Then using the estimated rotation angles, namely, the pitch angle φ_x , roll angle φ_y and azimuth angle φ_z , the rotation matrix $\mathbf{R}$ given in (3-5) is determined.

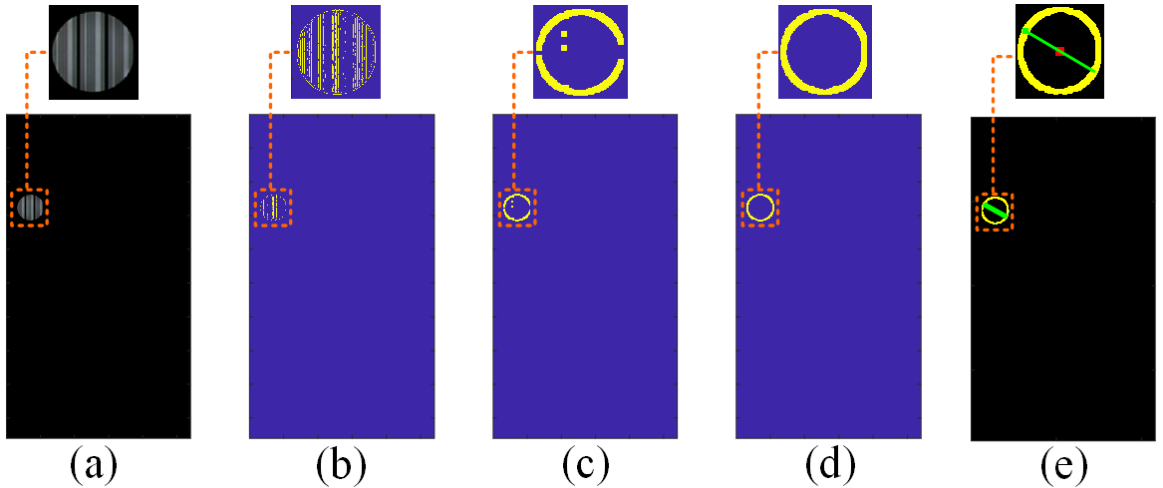


Figure 3.5: The image processing procedures of boundary detection: (a) original image, (b) binary image, (c) patterns eliminated, (d) detected boundary, (e) detected center in red and major axis in green.

To determine the geometric features of the LED image, the system has to detect the elliptical boundary curve in the captured image and the image processing procedures includes converting the original image to a binary image, deleting the rolling shutter patterns, detecting the boundary and estimating the ellipse center and major axis, as illustrated in Figure 3.5. The original RGB image, as shown in Figure 3.5(a), will firstly be converted to a grayscale image. Then the Sobel filter is applied to obtain a binary image, as shown in Figure 3.5(b). Since the rolling shutter patterns containing the unique ID of the captured LED are also detected, the pixels representing the patterns should be deleted by image processing steps, and the result is as given in Figure 3.5(c). As illustrated in Figure 3.5(c), a small number of the pixels on the rolling shutter patterns inside the elliptical boundary may not be deleted after the process. Additionally, the pixels on the boundary curve may be wrongly eliminated. Therefore, we further apply opening operation to exclude the pixels inside the LED boundary, and apply dilation operation to fill the boundary. Then, a complete LED boundary is obtained as illustrated in Figure 3.5(d).

After finding the elliptical boundary curve, the next step is to determine the geometric features of the ellipse. Here, we propose a center searching method to find the geometric features of the detected ellipse, namely, the pixel coordinates of the ellipse center and the lengths of the major and minor axes. For each pixel in the image, we compute the maximum distance between the pixel and the ellipse boundary and then find the shortest maximum distance after searching all the pixels in the picture. The one with the shortest maximum distance is the center of the ellipse and the shortest distance is the length of semi-major axis. The detection results of the center and the major axis are illustrated in Figure 3.5(e). To further determine the minor axis of the ellipse, we draw a line starting from the center and vertical to the major axis, and find the intersection point of the boundary and the line. Then, the distance between the intersection point and the center is the semi-minor axis.

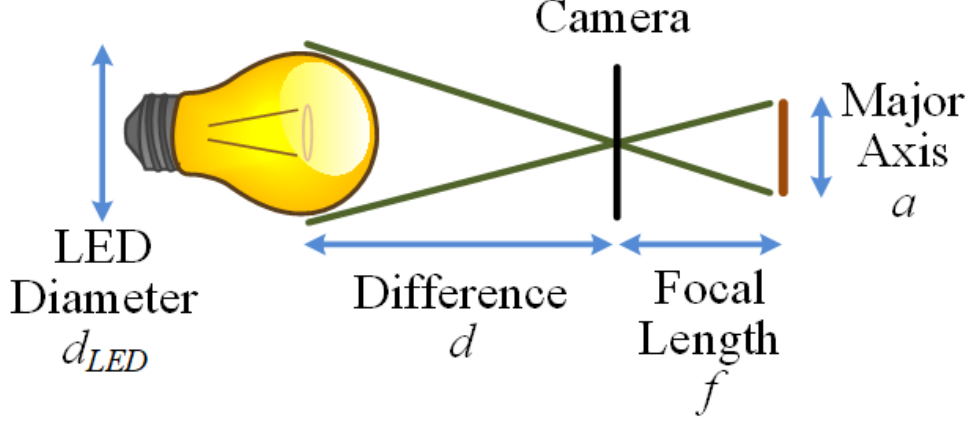


Figure 3.6: The pinhole camera model describing the relationship between the major axis length of the LED image and the distance between the LED and the camera.

The scaling factor z_c in (3-3) and the z-axial translation z_0 in (3-4) can be calculated using the pinhole camera model [24]. Since the size of the LED is much smaller than the distance between the LED and the camera, the projection can be supposed as a weak perspective projection [15]. It is assumed that the distance between the camera and every point on the LED is the same. According to the pinhole camera model as illustrated in Fig. 6, the distance between the camera and the LED d can be obtained by

$$d = \frac{d_{LED}}{a} f, \quad (3-6)$$

where d_{LED} is the diameter of the LED, f is the camera focal length and a is the major axis length of the detected ellipse. If we set the point on the ground directly below the LED as the origin of the VLP system, that is, the z-axial world coordinate of the LED Z_{WLED} is equal to the height of the LED, then the z-axial translation z_0 and the scaling factor z_c can be determined as

$$z_0 = d - Z_{WLED}, \quad (3-7)$$

$$z_c = Z_{WLED} - d. \quad (3-8)$$

3.4 Partially Blocked Led Image Compensation Method for VLP Systems

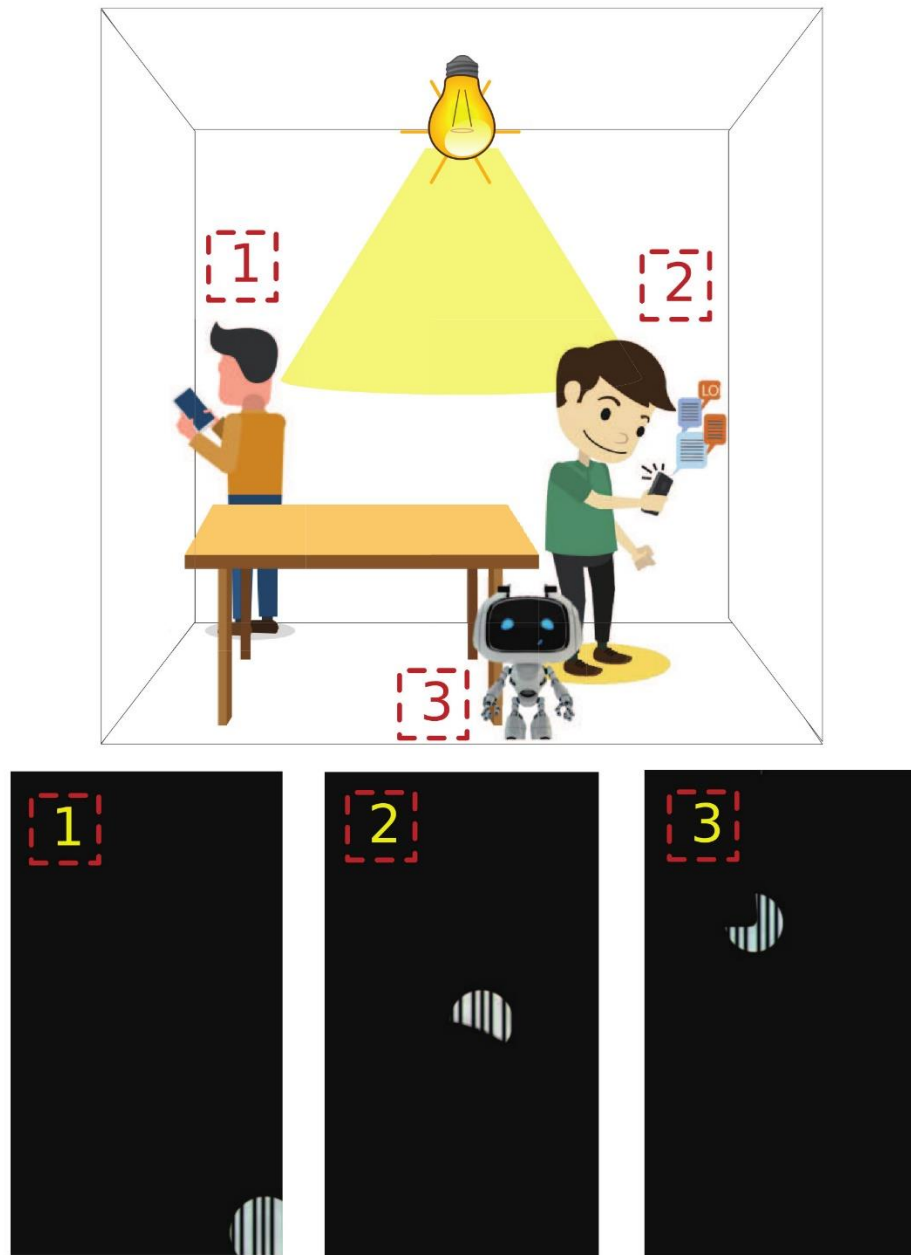


Figure 3.7: Incomplete LED images caused by a limited FOV, blockage of a human head and furniture.

As described in Section 3.1, in realistic scenarios the camera may capture incomplete LED images due to a limited FOV, blockage by a human head or furniture, as shown in Figure 3.7. Explicitly, when the camera is placed on the margin of the FOV, only part of the LED light can be captured. Additionally, when someone is standing below the LED lamp and holding a

smartphone to seek his location, his head partially blocks the LED light. Similarly, when a robot is controlled to track an object, it may move under tables or shelves, and the furniture will partially obstruct the LED light. Therefore, we propose to realize positioning using incomplete LED images and further enhance the robustness of the single-LED-based VLP system. For LED images with a shadow, it is noteworthy that after detecting the boundary of the LED image, a step to delete the points not on the ellipse should be added before finding the geometric features. Otherwise, the curve projected by the shadow will also be detected after the image processing methods, leading to additional errors in determining the geometric features. Fig. 8 illustrates the detected shadow curve when the LED light is blocked by a human head. Here we delete the points on the shadow curve with the aid of the property that the points wrapped by the shadow curve are black. The first step is to find the midpoint of every two pixels on the detected curve and then delete the two pixels if the binary value of the midpoint is 0. After checking every pair of points on the detected curve, only those on the LED elliptical boundary curve are kept.

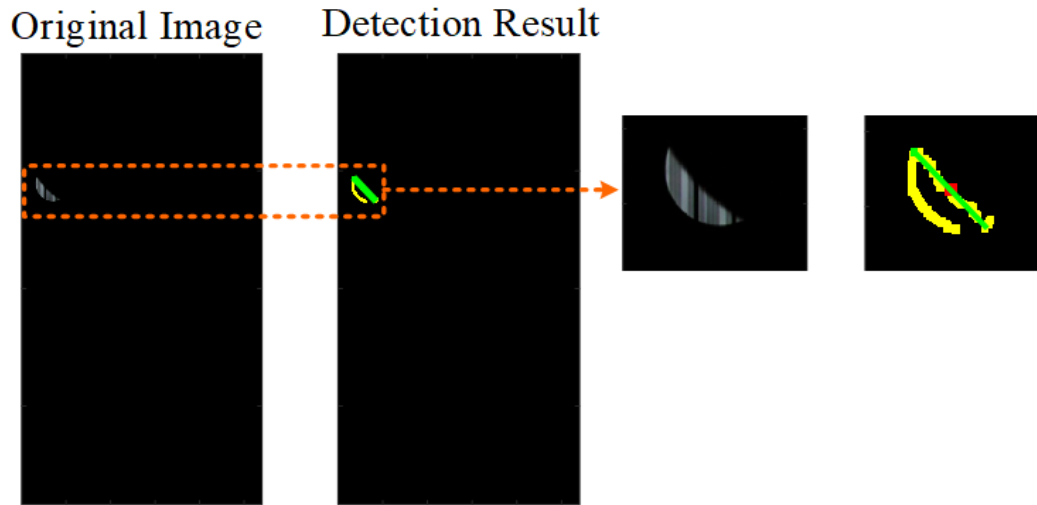


Figure 3.8: The detected shadow curve caused by human head blockage.

Since the center searching method described in Section 3.3.2 relies on finding the shortest maximum distance between all the pixels and the ellipse curve to determine the center, it may give poor performance when the major axis of the elliptical LED boundary is partially blocked by a shadow. Furthermore, Hough circle detection is widely used to find the radius and center [23]. However, the generalized Hough transform can only achieve high efficiency when a large

number of votes dropping in the correct elliptical boundary and then the right boundary can be directly detected, meaning that the elliptical boundary must contain enough points. Therefore, the efficiency of the generalized Hough transform relies on the quality and the number of the input points. Thus, we propose a boundary fitting method to compute the ellipse equation by at least five points on the incomplete elliptical boundary curve. The implicit equation of an ellipse is given by

$$Ax_e^2 + 2Bx_e y_e + Cy_e^2 + 2Dx_e + 2Ey_e - 1 = 0, \quad (3-9)$$

where (x_e, y_e) are the pixel coordinates of the points on the elliptical curve, A , B , C , D and E are the parameters in the implicit equation to be determined.

After solving for A , B , C , D and E , the pixel coordinates of the ellipse center (x_{ec}, y_{ec}) , the length of the major axis a and the length of the minor axis b can be determined by

$$x_{ec} = \frac{BE - CD}{AC - B^2}, \quad (3-10)$$

$$y_{ec} = \frac{BD - AE}{AC - B^2}, \quad (3-11)$$

$$a = 2 \sqrt{\frac{2(Ax_{ec}^2 + Cy_{ec}^2 + 2Bx_{ec}y_{ec} - 1)}{A + C + \sqrt{(A - C)^2 + 4B^2}}}, \quad (3-12)$$

$$b = 2 \sqrt{\frac{2(Ax_{ec}^2 + Cy_{ec}^2 + 2Bx_{ec}y_{ec} - 1)}{A + C - \sqrt{(A - C)^2 + 4B^2}}}. \quad (3-13)$$

3.5 Experiment and Evaluation

We build an experimental platform using a common circular LED luminaire and a Lenovo Android phone to verify the proposed VLP system, as shown in Figure 3.9. The experimental parameters and the camera options are summarized in Table 3.1. The experimental area is divided into a 50 cm $\times$ 50 cm grid size and 49 test points are selected. The smartphone is placed at these points and captures pictures. The tilted angle ranges about the x-, y- and z-axis are $-40^\circ \leq \varphi_x \leq 40^\circ$, $-40^\circ \leq \varphi_y \leq 40^\circ$ and $-60^\circ \leq \varphi_z \leq 60^\circ$, respectively. For each point, the smartphone captures four pictures with different tilt angles and different blocking ratios to avoid accidental errors and cover all the tilted angle range. Both positioning error and positioning error rate (PER) performances are measured for comparison. The PER describes the ratio of the positioning error to the distance given by [16]

$$PER = \frac{\sqrt{\Delta X_W^2 + \Delta Y_W^2 + \Delta Z_W^2}}{\sqrt{X_W^2 + Y_W^2 + Z_W^2}} * 100\%, \quad (3-14)$$

where (X_W, Y_W, Z_W) are the actual world coordinates of the camera and $(\Delta X_W, \Delta Y_W, \Delta Z_W)$ are the differences between the actual world coordinates and the estimated world coordinates using the proposed positioning methods.

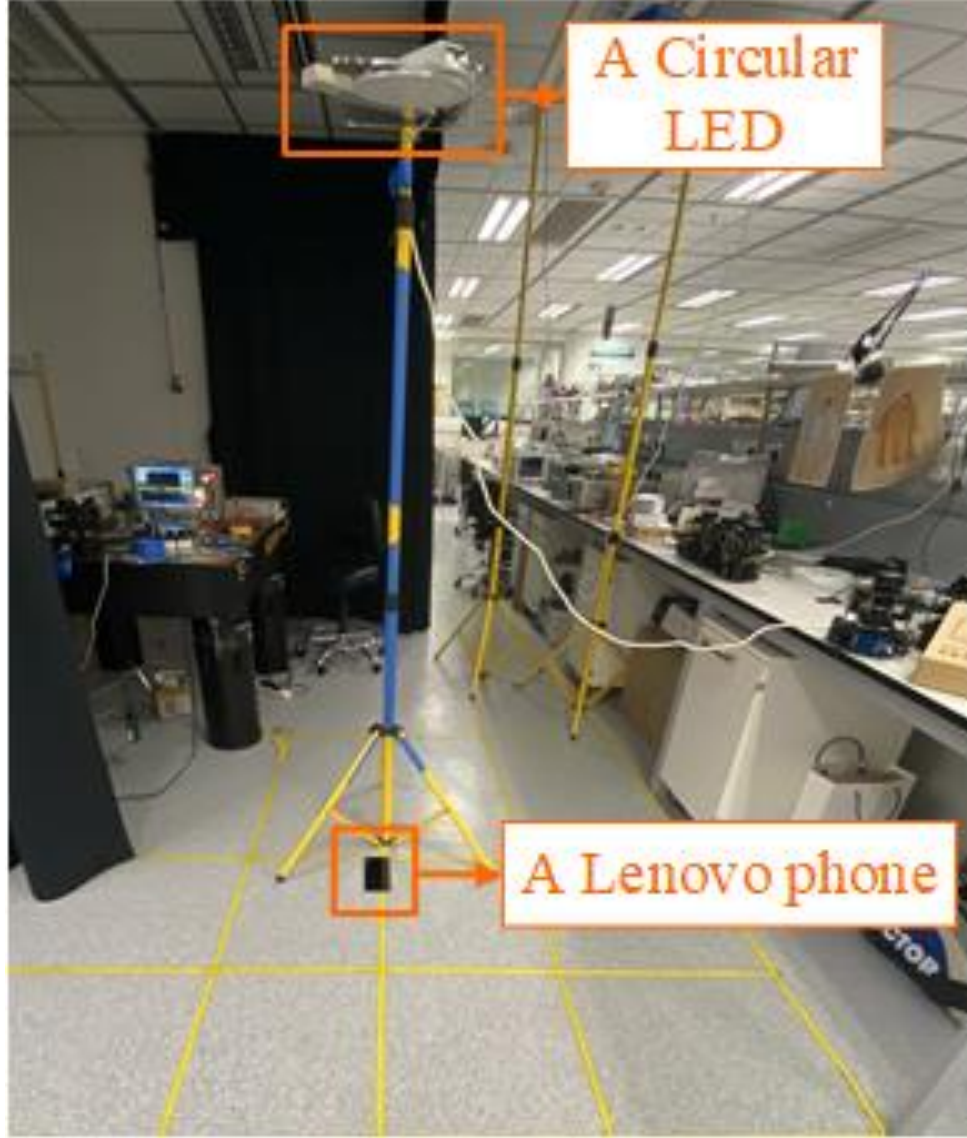


Figure 3.9: The experimental setup of the proposed single LED-based VLP system.

Table 3.1: Experimental Parameters

Horizontal Area	$3 \times 3 \text{ m}^2$
LED Height	1.86 m
LED Diameter	0.175 m
Phone Model	Lenovo PB2-690Y
Camera Resolution	1920×1080
ISO	100
Exposure Time	$\frac{1}{3000} \text{ s}$
Tilted Angle Range	$-40^\circ \leq \varphi_x \leq 40^\circ$
	$-40^\circ \leq \varphi_y \leq 40^\circ$
	$-60^\circ \leq \varphi_z \leq 60^\circ$

3.5.1 Performances of the Partially Blocked LED Image Compensation Methods

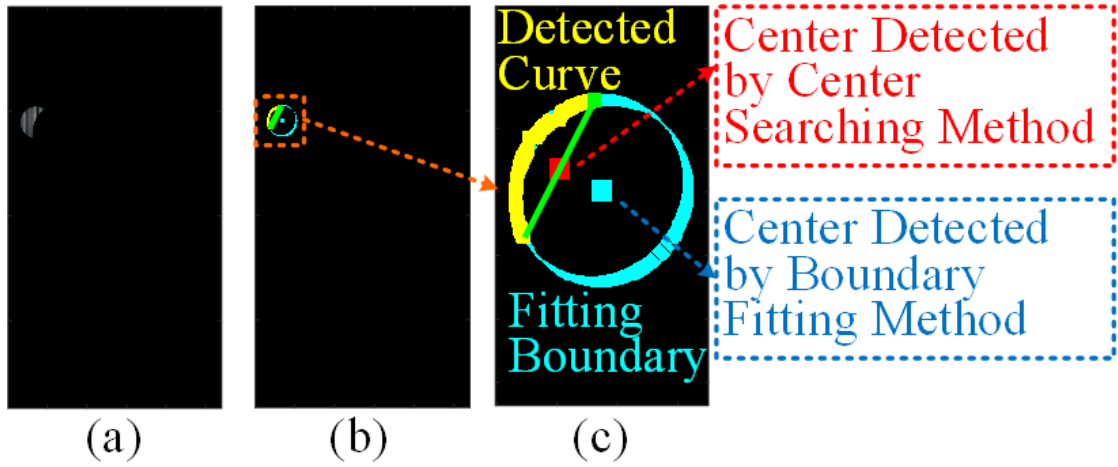


Figure 3.10: The detection results of center searching method and boundary fitting method: (a) the original captured LED image, (b) the detection results, (c) the zoomed in detection results.

Firstly, we compare the detection performances of the proposed center searching method and boundary fitting method. Figure 3.10 illustrates the estimated centers using the center searching method and boundary fitting method, respectively. The yellow curve is the detected curve after eliminating the shadow curve, showing that about 30% of the LED boundary is

captured. The red point is the ellipse center determined by the center searching method and the green line is the corresponding major axis. The blue point is the center estimated by the boundary fitting method and the blue curve is the fitting boundary. Since more than half of the LED image is blocked and the major axis of the elliptical boundary is not captured, the detection results of the boundary fitting method are more accurate than those of the center searching method. Therefore, the boundary fitting method can provide better robustness for the proposed positioning system.

3.5.2 Performances of the Tilted Receiver Camera Correction and Partially Blocked LED Image Compensation Methods for a Single-LED VLP System

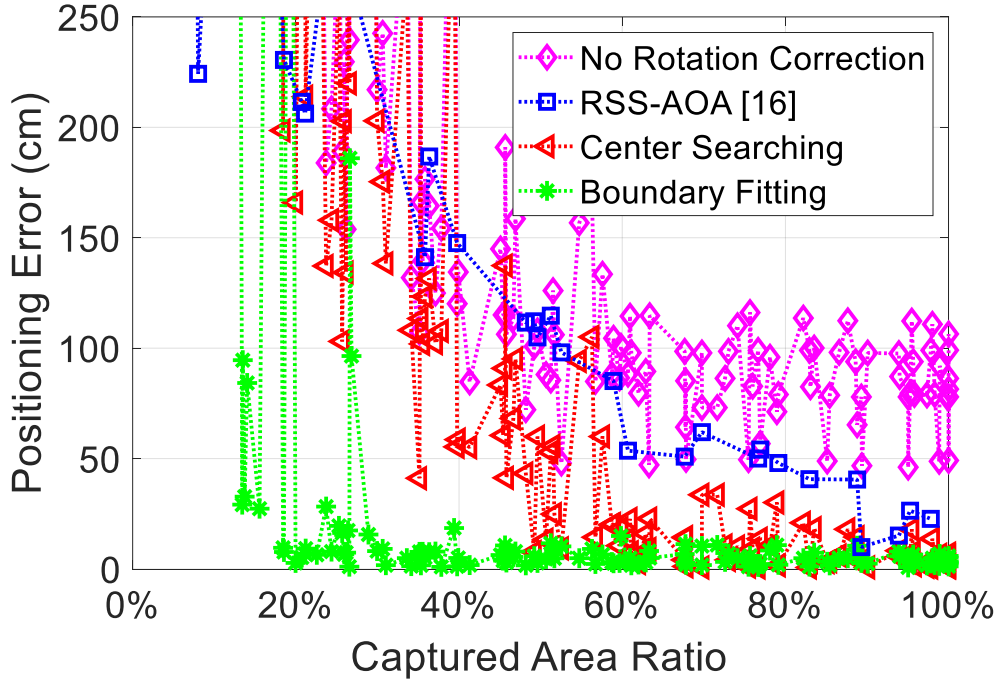


Figure 3.11: The comparison of the positioning error performance based on LED images with different captured area ratios among the two proposed methods, the RSS-AOA method [16] and positioning without rotation correction.

We measure the positioning error performances of the two proposed methods based on incomplete LED images and a tilted receiver in an experimental area of $2 \times 2 \text{ m}^2$ and compare the results with those of the received signal strength and AOA (RSS-AOA) method [16]. The positioning errors with different captured area ratios of the elliptical LED images are provided

in Figure 3.11. The captured area ratio is the ratio of the captured LED image area to the full LED image area. It can be seen from the figure that when no rotation correction is applied, and even when the full LED image can be captured, the positioning errors are higher than 50 cm. For the VLP methods with tilted receiver camera correction, the picture is mixed. When the captured area ratio is between 20% and 40%, that is, only about 20% to 40% of the LED image is shown, the positioning error of the proposed center searching method and RSS-AOA method [16] is generally higher than 100 cm. However, the positioning error of the proposed boundary fitting method is usually lower than 30 cm. This means that the proposed center searching method and the RSS-AOA method [16] are unable to realize positioning in this case, while the proposed boundary fitting method is able to provide coarse-precision positioning. When the captured area ratio is between 40% and 60%, the proposed boundary fitting method performs much better than the proposed center searching method and the RSS-AOA method [16], and when more than 60% of the LED image is captured, the performance of both the proposed center searching method and boundary fitting method is better than that of the RSS-AOA method [16]. When more than 90% of the LED image is captured, the average positioning errors of the center searching method and boundary fitting method can achieve 5.7 cm and 3.9 cm, respectively. As mentioned in Section 3.1, the average positioning errors of the circle geometry method [15], projective geometry method [17] and planes intersection-line method [18] are 17.52 cm, 11.2 cm and 5.58 cm, respectively. Therefore, in ideal conditions when a complete LED images is captured, the average positioning error of our proposed center searching method is comparable to existing image sensor-based VLP methods and the proposed boundary fitting method is better than the existing methods, whereas the two proposed methods outperform all the other systems in practical situations when the LED light is partially blocked or it is on the edge of a captured image.

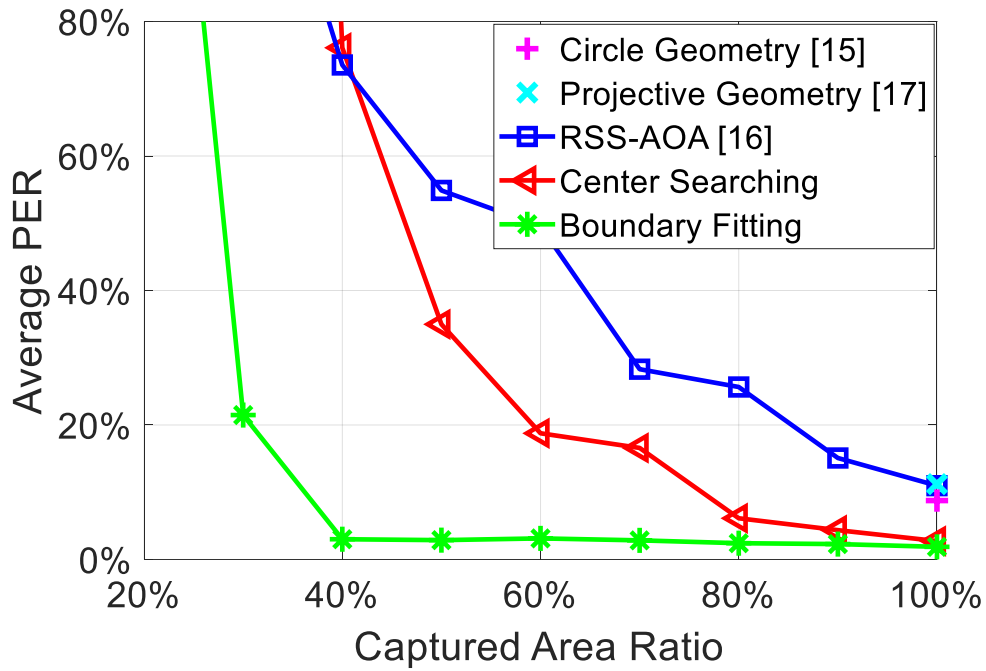


Figure 3.12: The comparison of average PERs based on LED images with different captured area ratios among the two proposed methods and the circle geometry method [15], RSS-AOA method [16] and projective geometry method [17].

Figure 3.12 illustrates the average PER at different intervals of the captured area ratio. As shown in the figure, when the captured area ratio is between 20% and 40%, compared with the proposed center searching method and the RSS-AOA method [16], the proposed boundary fitting method can achieve the optimal performance, with an average PER of about 20%. When the captured area ratio is higher than 40%, the average PER of the proposed boundary fitting method is lower than 4%. Both the proposed center searching and boundary fitting methods perform much better than the RSS-AOA method [16] when the captured area ratio is higher than 50%. Additionally, when the captured area ratio is higher than 80%, the two proposed methods can achieve an average PER lower than 6%. When a complete LED image is captured, the average PERs of the proposed center searching method and boundary fitting method are 2.73% and 1.87%, respectively, much lower than these of the circle geometry method [15], projective geometry method [17] and RSS-AOA method [16]. As indicated in [4], the positioning errors of VLP systems using LEDs are among 10-35 cm in experiments. Therefore, we aim to achieve an average positioning error lower than 10 cm. Since we verify our proposed

positioning methods in a $2 \times 2 \times 1.86 \text{ m}^3$ space, the expected average PER is aimed to be lower than 6%. Therefore, our two proposed methods can achieve comparable and even better positioning performance than existing approaches when a complete LED is captured, and surpass all the existing methods in combating the performance degradation caused by partially blocked LED light.

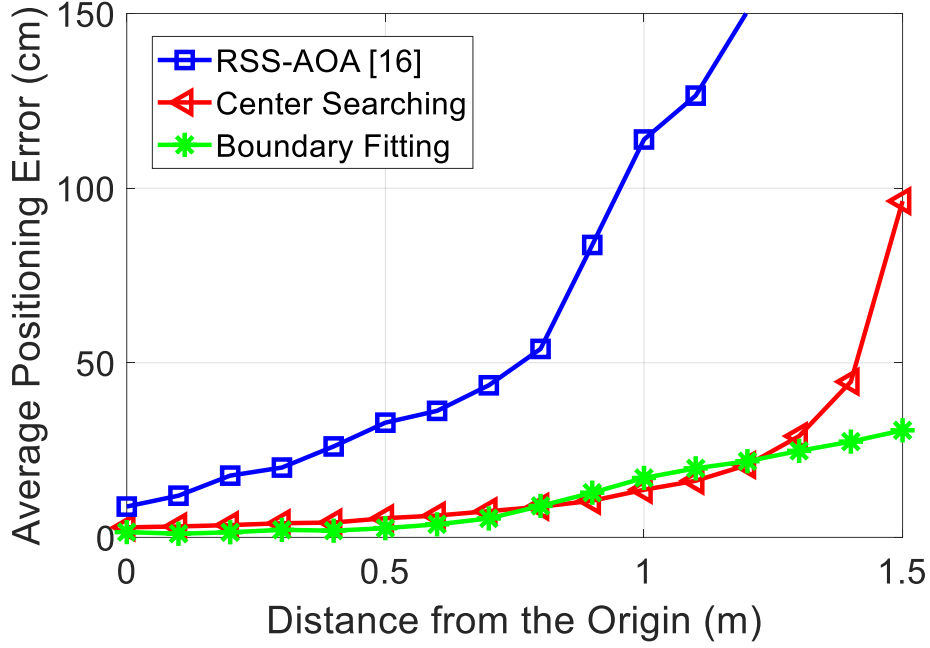


Figure 3.13: Coverage extension of the proposed center searching method and boundary fitting method along the margin of FOV.

Due to the limited FOV of the camera, if the smartphone is placed horizontally, the positioning area is limited to $2 \times 1 \text{ m}^2$ when the height of the LED is 1.86 m. Since the proposed methods can still achieve robust positioning when the camera is tilted about the x-, y- and z-axes and even when an incomplete LED image is captured, our positioning area can be extended to $3 \times 3 \text{ m}^2$ at the same height of the LED. Thus, we evaluate the positioning performances on coverage extension of the two proposed methods in an experimental scale of $3 \times 3 \times 1.86 \text{ m}^3$ and compare the results with the RSS-AOA method [16]. We fix the tilted angles of the smartphone as $\varphi_x = 0^\circ$, $\varphi_y = \pm 20^\circ$ and $\varphi_z = 0^\circ$, and move the smartphone along the y-axis in the world coordinate system with the interval of 0.1 m and take four pictures at the same point. The origin is set on the ground, directly below the LED. Figure 3.13 illustrates the average positioning error at different distances from the origin. When the distance

from the origin is shorter than 1.2 m, compared with the RSS-AOA method [16], both the proposed center searching method and boundary fitting method can provide better performance and effectively reduce the additional error caused by the tilted camera. When the distance from the origin is longer than 1.2 m, only incomplete LED images can be captured. In this scenario, the positioning error of the RSS-AOA method [16] goes much higher while the proposed boundary fitting method can still provide stable positioning service. When the distance is about 1.5 m, namely, on the margin of the area of $3 \times 3 \text{ m}^2$, about 20% of the LED can be captured, the proposed boundary fitting method can still reduce the negative effect of incomplete LED images and tilted camera. It is noteworthy that the proposed boundary fitting method has better coverage extension performance, and thereby has less tense requirement on the density of the LEDs in the venue.

3.6 Summary

In this work, we propose a tilted receiver camera correction and partially blocked LED image compensation method for indoor VLP systems. The proposed VLP methods do not require the LEDs to be placed at high density and can eliminate the additional positioning errors caused by a tilted receiver camera in realistic scenarios. We propose two methods to detect the geometric features of the captured LED images, and the experimental results show that both methods perform better than the existing single LED-based VLP methods in ideal scenarios. Furthermore, the two proposed methods outperform the existing methods and can effectively suppress the performance degradation when an incomplete LED image is captured. Therefore, the proposed methods can provide stable positioning services and further improve the robustness of the VLP system.

3.7 References

- [1] E. Kaplan and C. Hegarty, *Understanding GPS: Principles and Applications*. Artech House, 2005.

- [2] V. Cant'ón Paterna, A. Calveras Auge, J. Paradells Aspas, and M. A. Perez Bullones, "A bluetooth low energy indoor positioning system with channel diversity, weighted trilateration and Kalman filtering," *Sensors*, vol. 17, no. 12, pp. 2927–2927, 2017.
- [3] R. Zhou, X. Lu, P. Zhao, and J. Chen, "Device-free presence detection and localization with SVM and CSI fingerprinting," *IEEE Sens. J.*, vol. 17, no. 23, pp. 7990–7999, 2017.
- [4] 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. Tutor.*, vol. 20, no. 3, pp. 1963–1988, 2018.
- [5] D. Konings, N. Faulkner, F. Alam, E. M.-K. 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, 2019.
- [6] C. Yeh, C. W. Chow, H. Chen, J. Chen, and Y. Liu, "Adaptive 84.44- 190 Mbit/s phosphor-LED wireless communication utilizing no blue filter at practical transmission distance," *Opt. Express*, vol. 22, no. 8, pp. 9783–9788, 2014.
- [7] J. He, Z. Jiang, J. Shi, and Q. Tang, "An effective mapping scheme for visible light communication with smartphone camera," *IEEE Photon. Technol. Lett.*, vol. 32, no. 10, pp. 557–560, 2020.
- [8] M. K. Hasan, M. Z. Chowdhury, M. Shahjalal, M. M. Islam, and Y. M. Jang, "Optimum LED coverage utilization in OCC for effective communication with mobile robot," *J. Commun. Netw.*, vol. 22, no. 5, pp. 371–379, 2020.
- [9] B. Hussain, C. Lau, and C. P. Yue, "Li-Fi based secure programmable QR code (LiQR)," in *JSAP-OSA Joint Symposia*. Fukuoka, Japan: Optical Society of America, 2017, p. 6p A409 6.
- [10] C.-W. Chow, C.-Y. Chen, and S.-H. Chen, "Enhancement of signal performance in LED visible light communications using mobile phone camera," *IEEE Photon. J.*, vol. 7, no. 5, pp. 1–7, 2015.

- [11]G. Chen, W. Chen, Q. Yang, Z. Xu, L. Yang, J. Conradt, and A. Knoll, “A novel visible light positioning system with event-based neuromorphic vision sensor,” *IEEE Sens. J.*, vol. 20, no. 17, pp. 10 211–10 219, 2020.
- [12]R. Zhang, W.-D. Zhong, D. Wu, and K. Qian, “A novel sensor fusion based indoor visible light positioning system,” in *Proc. IEEE Globecom Workshops (GC Wkshps)*. Washington, DC, USA: IEEE, 2016, pp. 1–6.
- [13]P. Huynh and M. Yoo, “VLC-based positioning system for an indoor environment using an image sensor and an accelerometer sensor,” *Sensors*, vol. 16, no. 6, pp. 783–783, 2016.
- [14]T. Yuan, Y. Xu, Y. Wang, P. Han, and J. Chen, “A tilt receiver correction method for visible light positioning using machine learning method,” *IEEE Photon. J.*, vol. 10, no. 6, pp. 1–12, 2018.
- [15]R. Zhang, W.-D. Zhong, Q. Kemao, and S. Zhang, “A single LED positioning system based on circle projection,” *IEEE Photon. J.*, vol. 9, no. 4, pp. 1–9, 2017.
- [16]J.-W. Lee, S.-J. Kim, and S.-K. Han, “3D visible light indoor positioning by bokeh based optical intensity measurement in smartphone camera,” *IEEE Access*, vol. 7, pp. 91 399–91 406, 2019.
- [17]J. Hao, J. Chen, and R. Wang, “Visible light positioning using a single LED luminaire,” *IEEE Photon. J.*, vol. 11, no. 5, pp. 1–13, 2019.
- [18]H. Cheng, C. Xiao, Y. Ji, J. Ni, and T. Wang, “A single LED visible light positioning system based on geometric features and CMOS camera,” *IEEE Photon. Technol. Lett.*, vol. 32, no. 17, pp. 1097–1100, 2020.
- [19]Z. Xie, W. Guan, J. Zheng, X. Zhang, S. Chen, and B. Chen, “A high-precision, real-time, and robust indoor visible light positioning method based on mean shift algorithm and unscented Kalman filter,” *Sensors*, vol. 19, no. 5, pp. 1094–1094, 2019.
- [20]W. Guan, Z. Liu, S. Wen, H. Xie, and X. Zhang, “Visible light dynamic positioning method using improved Camshift-Kalman algorithm,” *IEEE Photon. J.*, vol. 11, no. 6, pp. 1–22, 2019.

- [21]Z. Zhang, “A flexible new technique for camera calibration,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1330–1334, 2000.
- [22]T. Ozyagcilar, “Implementing a tilt-compensated eCompass using accelerometer and magnetometer sensors,” *Freescale semiconductor*, AN, vol. 4248, 2012.
- [23]Y.-S. Kuo, P. Pannuto, K.-J. Hsiao, and P. Dutta, “Luxapose: Indoor positioning with mobile phones and visible light,” in *Proc. ACM MobiCom*, Maui Hawaii, USA, 2014, pp. 447–458.
- [24]W.-L. Du and X.-L. Tian, “An automatic image registration evaluation model on dense feature points by pinhole camera simulation,” in *Proc. IEEE Int. Conf. Image Process. (ICIP)*. Beijing, China: IEEE, 2017, pp. 2259–2263.

CHAPTER 4 High-Accuracy Indoor VLP System for Robots

4.1 Indoor Positioning for Robots

With the high-speed evolution of wireless technologies and the high demand of mobile devices, indoor positioning technology has enjoyed expansive development prospects. In the interior of buildings, densely populated cities and underground environments, traditional outdoor positioning technologies such as Global Positioning System (GPS) have poor signal coverage and large positioning errors [1]. Moreover, compared with other radio wave positioning technologies such as WLAN (wireless LAN/Wi-Fi), radio frequency identification (RFID), Bluetooth, VLC-enabled positioning technology can theoretically provide higher positioning accuracy and does not produce any electromagnetic interference [2]. Visible light positioning (VLP) systems compile the transmitted position information into a modulated signal and modulate it to the driving current of light emitting diodes (LEDs). When the positioning terminal enters the illumination area, it will receive and recognize the optical signals transmitted by the LEDs through a sensor such as a photodiode (PD) or an image sensor, and resolve the unique identification (ID) information of the LEDs. Then the corresponding location information is determined in the map database.

With the rapid development of artificial intelligence (AI) technology in recent years, the impact on the intelligent construction industry has continued to increase, and it has brought considerable changes in many subdivisions. Industry leaders represented by Country Garden and Evergrande have deployed a variety of robotic equipment in the new generation of intelligent construction projects, such as handling building materials, leveling the ground, building positioning, and so on. However, its level of intelligence is not very high. Although there is a huge demand for full-cycle inspection and quality assessment of indoor building quality on the market, there is no product solution that integrates autonomy, sampling, reconstruction modeling, and data analysis. This is also the point of rapid growth of the industry in the next 5-10 year. Considering the cost-effectiveness ratio, the model of using autonomous robots plus intelligent sensing equipment will be more than 5 times higher than that of

traditional manual labor, and is expected to drive tens of billions of market demand, so the research and implementation of this system is of great significance. Facing the urgent needs of the future development of the intelligent construction industry, intelligent robot equipment and intelligent detection technology have gradually become the most core basic reserve. The degree of system-level intelligence will directly determine the quality and efficiency of task completion, and indirectly reflect the difference in production value. Therefore, in order to solve the bottleneck caused by the above problems to industry applications, this project proposes a fusion integrated solution based on a mobile robot platform and VLC co-location technology. VLC technology is used to improve the positioning ability of the mobile robot platform, so that it has accurate autonomous operation capabilities.

4.2 Related Works

4.2.1 Position Estimation of Robot

In a robot positioning system, there are two types of sensors [3]. One is onboard sensors, which adhere to the robot body, such as the odometer and IMU. These sensors measure the robot's linear and angular velocities and accelerations with a high updating rate, and predict its position and orientation by previous measurement. An indoor positioning system based on wheel odometry is proposed in [4] by fusing the readings from an encoder, gyroscope, and magnetometer using a self-tuning Kalman filter coupled with a gross error recognizer. IMU-based inertial navigation is an important positioning technique in robot localization as it works autonomously. They present the benefit of high short-term accuracy and great anti-interference ability. However, inertial navigation has a major drawback of not able to provide long-term accurate positioning because of the cumulative positioning error increases over time. In [5], two IMUs are used to estimate the position, and the positioning performance is improved by the complementation of the relative relationship. The average positioning accuracy is lower than 20 cm over short periods of time. However, since the onboard sensors are subject to time-dependent integral error that increases over time [6], the accumulated error is still inescapable,

leading to the degradation of positioning accuracy. It will also reduce the level of the autonomy in various industrial tasks. Therefore, a high-robustness positioning system is needed.

The other type is external sensors, which are separated from the robot body, such as image sensors and light detection and ranging (LiDAR). These sensors are capable to of measuring an absolute position with the aid of a fixed global reference in the environment. In [7], a radio-frequency identification (RFID) reader is mounted on the robot to track its position in the scene where RFID tags are placed at each intersection of structured environment ways. A Bayesian filter-based robot positioning system with RFID tag collecting is proposed in [8], and the average positioning accuracy is about 50 cm. In [9], the robot is equipped with an array of microphone, and the positioning is achieved using time difference of arrival (TDOA). An unscented Kalman filter-based position estimation method is proposed in [10], where a tachometer is mounted on the robot. To increase the positioning accuracy, more sensors, such as IMUs, are needed. In [11], a biomimetic radar sensor-based positioning system is proposed, and it can locate a robot with an average accuracy of 35 cm. Another robotic positioning scheme was based on simultaneous localization and mapping (SLAM) and laser sensors, and was presented in [12] using Monte Carlo localization and convolutional neural network (CNN) algorithm. However, the average accuracy is limited 40 cm which is insufficient for robotic applications.

VLC is a powerful technology for future generations of mobile network well beyond 5G. Based on LED and VLC technologies, visible light positioning (VLP) can use the LED lightings to transmit position information. Compared with the above works, VLP can achieve much better, centimeter-level, positioning accuracy.

According to receiving sensors, VLP technology can be divided into PD-based positioning and image sensor-based positioning. In [13], the authors proposed and demonstrated a PD-based VLP system with machine learning technique applied to enhance positioning accuracy. In image sensor-based VLP systems, image-processing approach is applied at the receivers to convert the received 2D image into a three-dimensional image. Compared with PD-based VLP technology, image sensor-based VLP technology is less affected by ambient light, and can

realize higher positioning accuracy. Moreover, image sensors can be combined with smart mobile devices such as smartphones to truly implement VLP technology from research lab to commercial applications.

The consideration of positioning algorithm includes positioning accuracy, real-time and robustness [14]. However, most existing studies on image sensor-based VLP can only achieve static positioning and aim at enhancing the positioning accuracy. The method can suppress the positioning errors caused by rotation and enhance the robustness of image sensor-based VLP systems. Nowadays, there is an increasing demand for robots to conduct more challenging and smart tasks, and at the same time the operation environments of robots has become more complicated. Indoor complex scenarios, such as domestic or industrial workspaces, contain various blind corners that are hard to detect. These kinds of application and environmental status will actually influence the positioning accuracy of indoor robotic platforms. Furthermore, such systems are very demanding and challenging since all the above situations and types of movement are characterized by high complexity and diversity. Two requisites of indoor robot positioning systems are real time ability and high accuracy, which will improve the efficiency of robot work. In [21], a loosely-coupled VLP-inertial fusion method was proposed, to improve positioning robustness under LED shortage/outage with an inertial measurement unit (IMU) and rolling shutter camera. However, it is challenging for a robot to locate itself in a complex and moving scenario only with a single sensor. Therefore, the localization and navigation of the robot is a challenging problem, and VLP should be the most suitable solution for indoor mobile robot positioning.

4.2.2 Path Planning Algorithms

In a navigation process, path planning and kinematics control is applied to the robot to reach a required position. Path planning includes a global planner and a local planner. The objective of global path planning is to build a feasible path from the starting point of the robot to the goal set by the controller. Global path planning is based on the static map layer and is updated with a relatively low frequency, assuming complete knowledge of the obstacles is obtained. By contrast, local path planning mainly follows the global path and at the same time makes the

robot avoid the dynamic obstacles detected by the sensors. Therefore, local path planning is based on the obstacle map layer determined by the data from the sensors on the robot and is updated with a relatively high frequency, assuming there is still unknown obstacles in the area to be perceived.

Table 4.1: Comparison between global path planning and local path planning.

Global Path Planning	Local Path Planning
Static map-based	Sensor-based
Comparatively low update frequency	Comparatively high update frequency
Assume perfect knowledge of the area to be perceived (static map)	Assume imperfect knowledge of the area to be perceived
Determine a beneficial path to the set goal	Follow the global path while avoiding obstacles

Dijkstra's algorithm is first proposed by a computer scientist called Edsger W. Dijkstra, and becomes one of the most universal global path planning algorithms [15]. It is capable of determining the shortest path from one vertex to other vertices. The main idea of Dijkstra's algorithm is that it starts from the starting point and adopts the strategy of greedy algorithm. It regards the starting point as the center point and expands outwards layer by layer until it reaches the destination [16]. As Dijkstra's algorithm uses greedy algorithm, the number of nodes on the graph determines its speed of computation. The advantage of Dijkstra's algorithm is that it can solve the optimum shortest path, but it takes longer computation time to search for the shortest path. A* algorithm is another widely used global path planning algorithm, and is first proposed by Peter Hart, Nils Nilsson and Bertram Raphael from Stanford Research Institute [17]. It is a heuristic search algorithm. In other words, it is an algorithm to find the optimal solution using heuristic function in a finite solution space that can be exhausted. A* algorithm can find the optimal solution, but highly relies on the heuristic function and has high computational complexity.

For local path planning, elastic band (EBand) planner has two essential components: contraction force and repulsion force [18]. Contraction force is used to generate the shortest

path between the start point and goal point, and repulsion force is aimed to avoid the path from the obstacles. EBand local planner can determine the shortest path. However, it does not consider the robot's kinematics and thereby is sensitive to the acceleration related parameters. Timed elastic band (TEB) local planner is an extension of EBand planner by decreasing the trajectory execution time [19]. It can bypass obstacles even when the obstacles are right in front of the robot. Therefore, it has good performance in dynamic-obstacle avoidance. However, it requires high computational complexity. Dynamic Window Approach (DWA) local planner takes robot's kinematics into consideration and requires relatively low computing power. The first step of DWA local planner is to discretely samples robot velocity including linear and angular velocities. Then forward simulation is performed to evaluate each trajectory in terms of calculating a cost function. The cost function is computed by the sum of the distance between the planned path and the endpoint of the trajectory, the distance between the local goal and the endpoint of the trajectory and the maximum obstacle cost along the trajectory. By comparing the cost function of each trajectory, the best trajectory is determined and the corresponding velocities are sent to the robot base. The disadvantage of DWA local planner is that it has poor performance in highly complex and dynamic-obstacle environment.

4.3 VLP-based Mobile Robot Platform

In this chapter, we propose an image sensor-based indoor VLP demonstration based on the studies on the identification of the LED-ID position information, positioning accuracy, real-time and robustness. Figure 4.1 describes the hardware and processing flow of the proposed robotic localization platform. The main contributions are as follows:

1. Design a smart LED system aimed for VLC positioning. VLC modulated LEDs are used as location beacons to provide location signals that are captured by the camera mounted on the robot to calculate the three-dimensional position of the robot with cm-level accuracy. Each smart LED generates a unique ID with a wireless control module. Bluetooth Low Energy (BLE) SoC is equipped to support iBeacon and VLC data/frequency and implement a wireless control system for these LEDs. The VLC signal adopts OOK modulation scheme by using designed data pattern to control the on and off states of a power switch.

2. Design a VLC high-precision three-dimensional imaging positioning algorithm based on a single LED. As VLP systems need to consider the accuracy and real-time problem at the same time, this demonstration proposes a high-precision VLC three-dimensional imaging localization algorithm based on a single LED.

3. Based on the above algorithms, we build a VLP-based mobile robot experiment platform. Smart LEDs are used for VLC function. The camera mounted on the robot will capture the images of LEDs, use ID recognition algorithm to identify the IDs and then get the position with image-based algorithm.

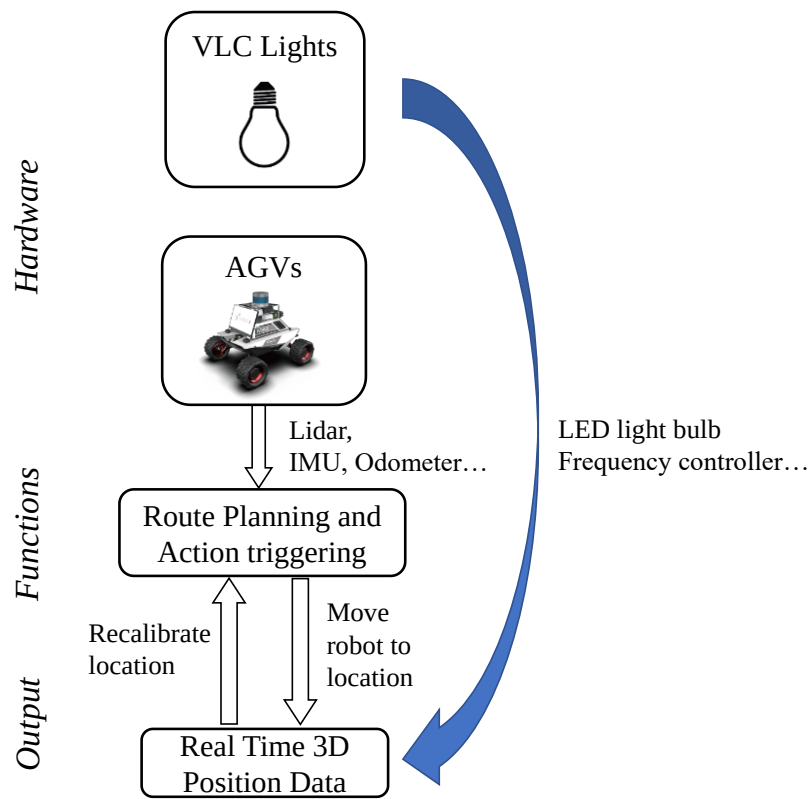


Figure 4.1: Diagram illustrating VLC sources used for robotic 3D localization.

4.3.1 Smart LED System for VLP

In the implementation of this demonstration, the whole system consists of two parts: intelligent lighting and image sensor-based VLC light tracking. Figure 4.2 shows our VLP platform for a single robot navigation system deployed at our lab (Integrated Circuit Design Center, 3/F, CYT Building, HKUST). As for the intelligent lighting, we use one smart LED as the signal source of VLC. To easy install and deploy the transmitter, we use a universal VLC

modulator to convert existing LED lighting into VLC modulated smart lighting for indoor positioning [22]. The unique identifiers (UIDs) provided by the LEDs are set as the critical points and the UIDs are stored in the memory (flash) integrated in BLE SoC [23]. BLE SoC is the core chip of wireless control module. With the embedded BLE, the lights also support geofencing. To avoid from interruption, we use common serial peripheral interface (SPI) with direct memory access (DMA) to control the generation of VLC data in BLE SoC. The frequency of SPI changes along with the sequence of data while DMA works in a repeated mode. Then all of the stored data are loaded one-off from flash to RAM and used to generate the control signal to modulate the light of those LEDs by adopting OOK modulation scheme. Then the data of UIDs can be transmitted without intervention. After the data sequence comes out from the previously determined I/O pin, it will be pushed out to the gate of NMOS power switch then control the on and off states. The LEDs are connected in series of the NMOS power switch and they will repeatedly turn on and turn off following the VLC data pattern. The UIDs broadcast by LEDs can be detected with the camera working at rolling shutter mode as captured images. At the receiver side, image processing and VLC decoding algorithm is applied to recover the embedded digital code from the captured patterns [24].

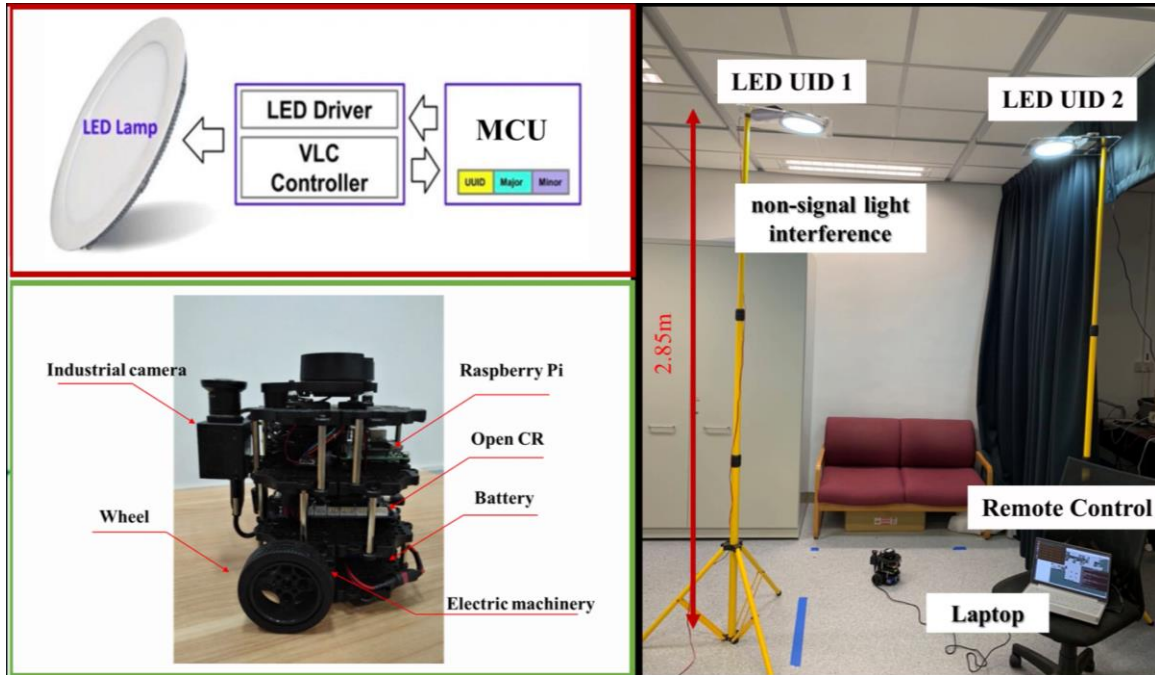
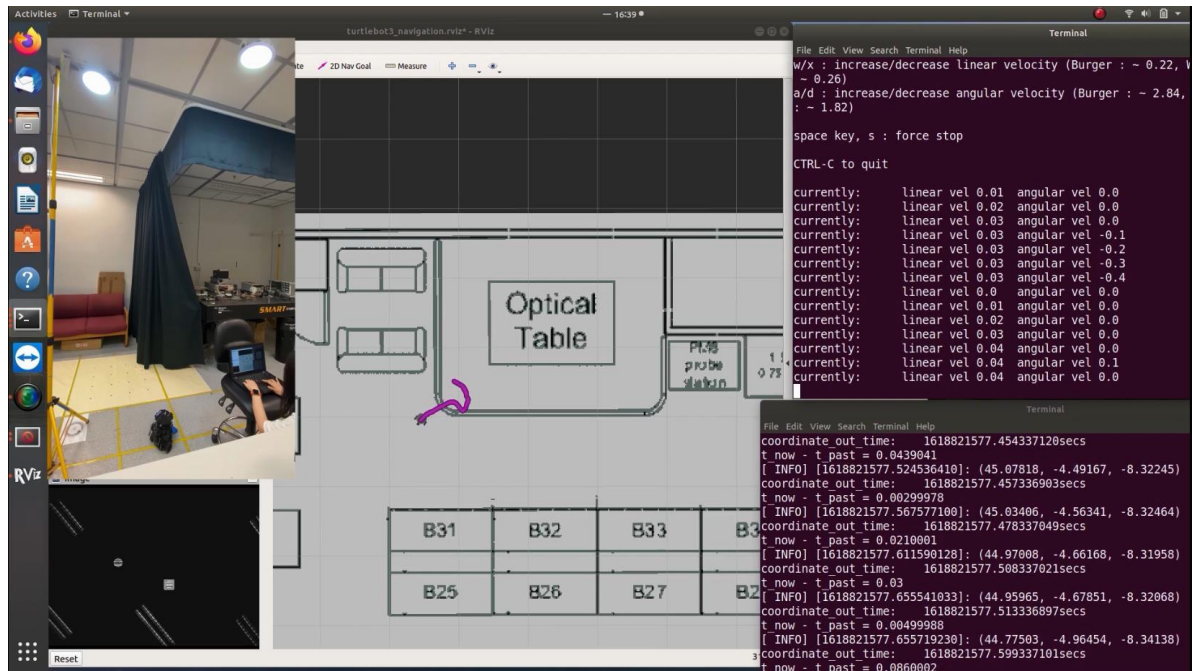


Figure 4.2: Demonstration setup of high-precision positioning system based on VLC smart lighting.

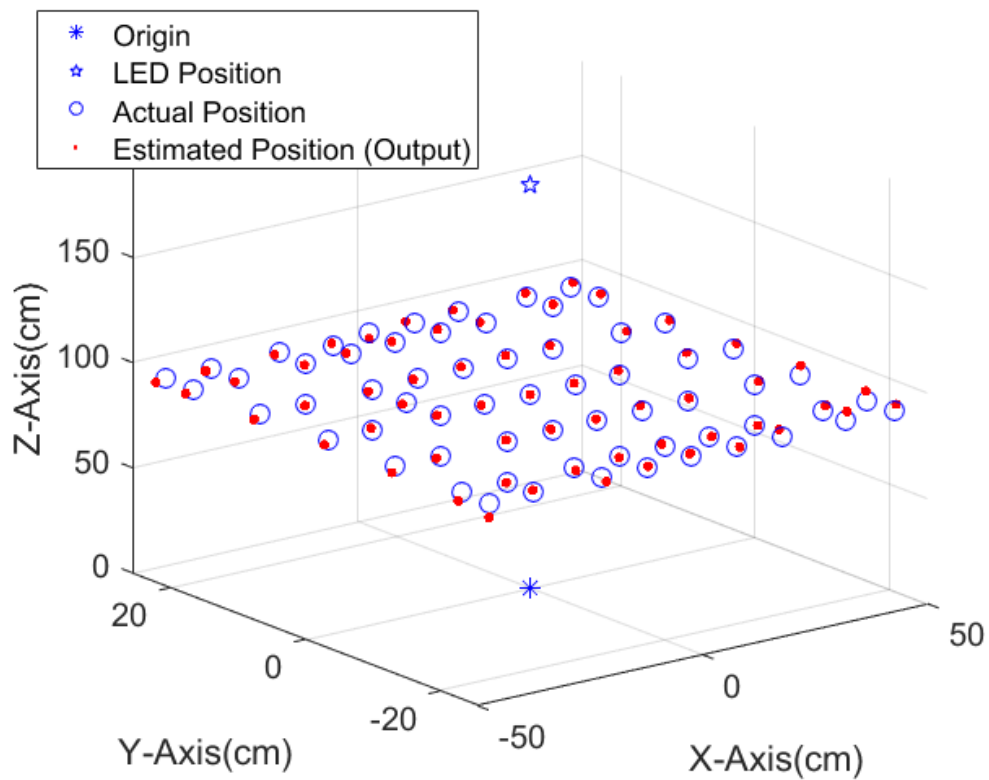
In this demonstration, we utilized LED panels with the diameter of 17.5cm and they are mounted on poles with the height of 2.85 m. The BLE SoC we used for wireless control module is DA14580_QFN40. For the receiver end, we use a camera mounted on the top of the robot to capture the images of the LEDs and use remote control laptop to process those images and dealing with tracking algorithm which is the second part of this system.

4.3.2 Real-time Robotic Localization System

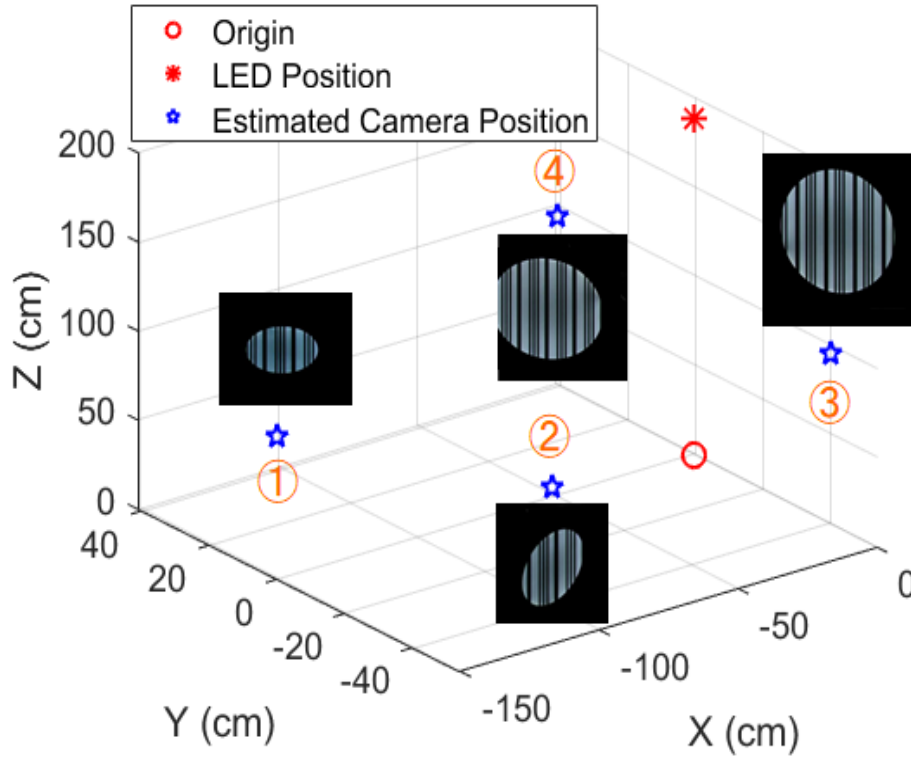
In order to realize the image sensor-based VLC light tracking function, the first task is to obtain the LED-ROI region by using the VLC dynamic positioning tracking detection algorithm. Secondly, we need to identify the ID position information with the LED-ID recognition algorithm. Image processing technique is performed to extract the images, based on which the LED-ID position information is identified. By pre-establishing the LED-ID database, the machine-learning algorithm is used to perform feature matching on the LED-ID light stripe code to realize LED-ID identification. Lastly, the LED image-based localization algorithm is applied for VLP. After accurately identifying the LED-ID, the position coordinate of the LED is obtained, the position of the terminal relative to the LED in the locating position area can be obtained by the imaging-positioning algorithm, thereby achieving indoor positioning [25]. For the image processing (LED-ID feature extraction) procedure, the VLC tracking detection algorithm is applied and relative coordinate of the camera to the LED is calculated. The Turtlebot3 robot kit is based on a Raspberry Pi 3B (Quad ARM Cortex-A53 Core 1.2 GHz Broadcom BCM2837 64 bits CPU and 1 GB RAM). The operating system of the Turtlebot3 robot is Ubuntu 16.04 MATE, and the operating system of the laptop is Ubuntu 18.04 LTS. The ROS release is ROS Melodic.



(a)



(b)



(c)

Figure 4.3: Experiment results of the proposed robot positioning system: (a) experiment setup, (b) VLP results when the camera moves on a plane, (c) VLP results when the camera rotates and is partially blocked in 3D space.

Figure 4.3 shows the experiment results of the proposed real-time robotic positioning system. The gray area in the middle of Figure 4.3(a) is the digital map of the perceiving area (Integrated Circuit Design Center, 3/F, CYT Building, HKUST). The gray block logo on the map is the real-time position of the robot. The real-time view of the camera mounted on the robot facing to the ceiling is shown at left bottom of Figure 4.3(a). The purple dots on the digital map are the positioning results of moving trace of the robot, which is controlled by the remote laptop. The terminal on the right top of Figure 4.3(a) shows the linear and angular velocities of the robot which are controlled by the keyboard. The terminal on the right bottom of Figure 4.3(a) gives the real-time 3D position of the robot estimated by the proposed single LED-based VLP system. Figure 4.3(b) shows the positioning results when the camera moves on a plane. The altitude difference between the plane and the LED is about 0.8 m and the average positioning

error is about 2 cm. Figure 4.3(c) illustrates the 3D VLP results when the camera has 3D rotation and is partially blocked in 3D space.

4.4 Application of Robot Positioning and Navigation System: Panorama Creation

Compared with ordinary pictures, panoramas are more vivid and interactive. Using panoramas to replace ordinary pictures is a rising publicity method in recent years. Panoramas give an all-round display instead of a single perspective and have been widely used in many places such as real estates, tourisms, hotels, and KTVs. Panorama can be created by human holding a 360-degree camera as shown in Figure 4.4. However, human taking pictures with hand-held cameras may cause camera shake. Furthermore, a 360-degree image has severe image distortion as shown in Figure 4.4(b). Therefore, we propose to use a robot mounted with an ordinary USB camera to take images and generate a panorama by rotating at a target point. By using the proposed robotic positioning and navigation system, the robot can be sent to a given location with high accuracy and obtain complete information of the environment. Furthermore, compared with a 360-degree camera, a USB camera is much cheaper.

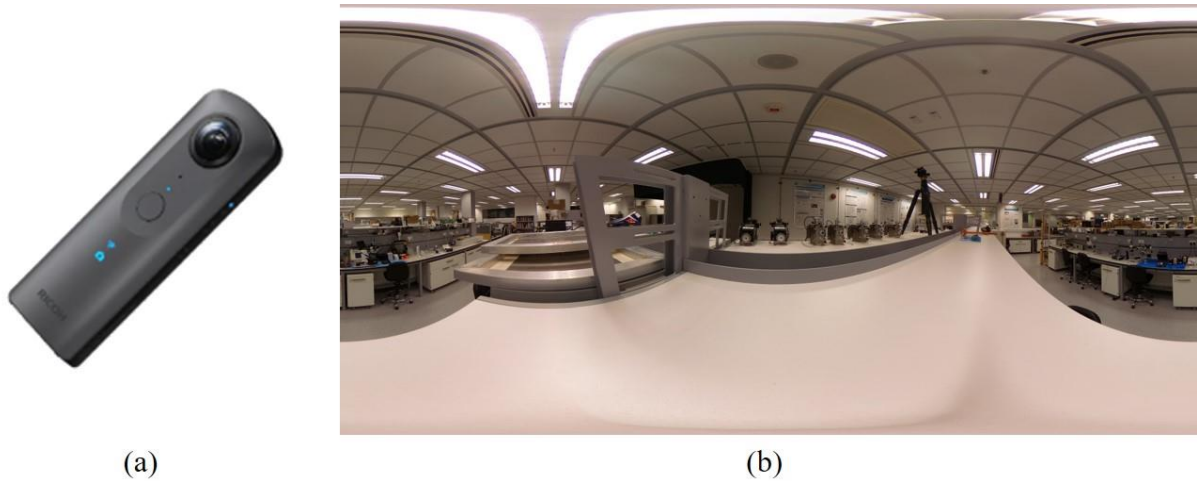


Figure 4.4: 360-degree image capture: (a) a 360-degree camera, (b) a typical 360-degree image

4.4.1 Real-time Robotic Navigation System

In ROS platform, navigation is achieved by a 2D navigation stack reading in the information from odometer, sensors and a goal pose sent by the controller. The map layer in ROS is

composed of a static map layer, an obstacle map layer and an inflation layer. The static map layer is usually based on a premade digital map, such as a sensor map or a floor plan. The obstacle map layer is constructed using the dynamic obstacle information detected by the 2D/3D LiDAR on the robot. The inflation map is the expansion of the dynamic obstacles on the obstacle map and the static obstacles on the static map.

Furthermore, using the obstacle information, a cost map is built during navigation process. Based on the cost map, a global and local path plan is developed. ‘move_base’ is the main package in the navigation stack and consists of five main nodes given by ‘global_planner’, ‘global_costmap’, ‘local_planner’, ‘local_costmap’ and ‘recovery_behaviour’. In our experimental setup for robot navigation, we use Dijkstra’s algorithm based global planner and DWA local planner.

4.4.2 Image Taking at a Target Point

In the proposed VLP-based image taking robot platform, a USB camera is mounted on the top of the robot as shown in Figure 4.5. The USB camera is Phottix PC-20 FHD Webcam and the resolution of the captured images are set as 640×480. Other setup is the same as that described in Section 4.3. Figure 4.6 gives the diagram of the proposed VLP-based image taking robot system. To navigate a robot mounted with a USB camera to a target point to take images and generate panoramas, the first step is to send the exact location of the target point to the robot in terms of 3D coordinates of the point or a robot pose on RViz. When receiving the navigation task, the robot will move to the goal under the guidance of Dijkstra’s global planner and DWA local planner. After reaching the goal, the robot will start rotating automatically, capture images with the USB camera and rename the images with the location information obtained from the VLP system. When rotating for 360 degree, the robot will stop automatically and wait for the next goal.



Figure 4.5: The robot equipped with a typical USB camera in the proposed VLP-based image taking robot system.

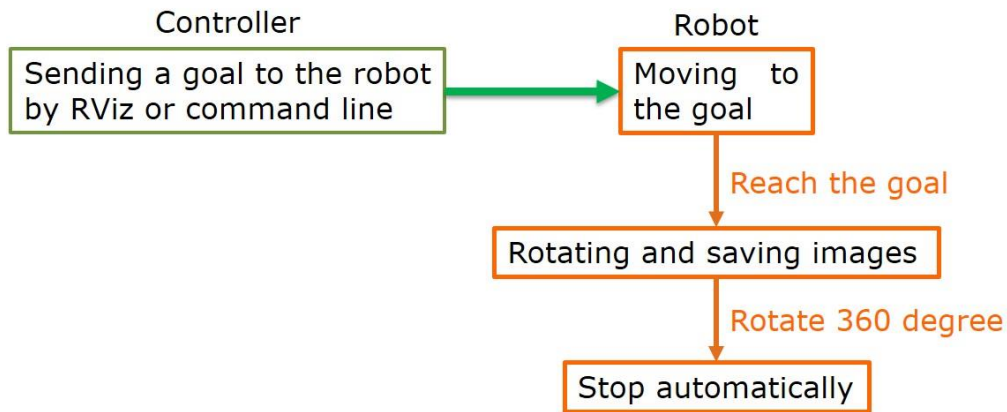
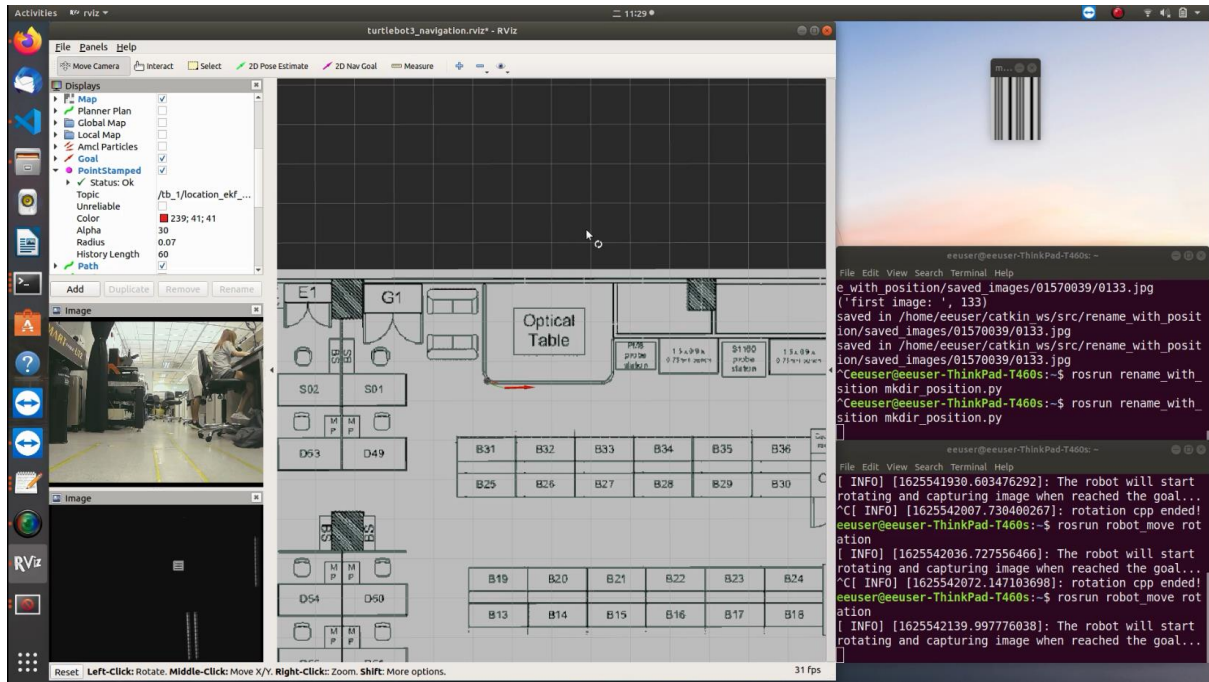


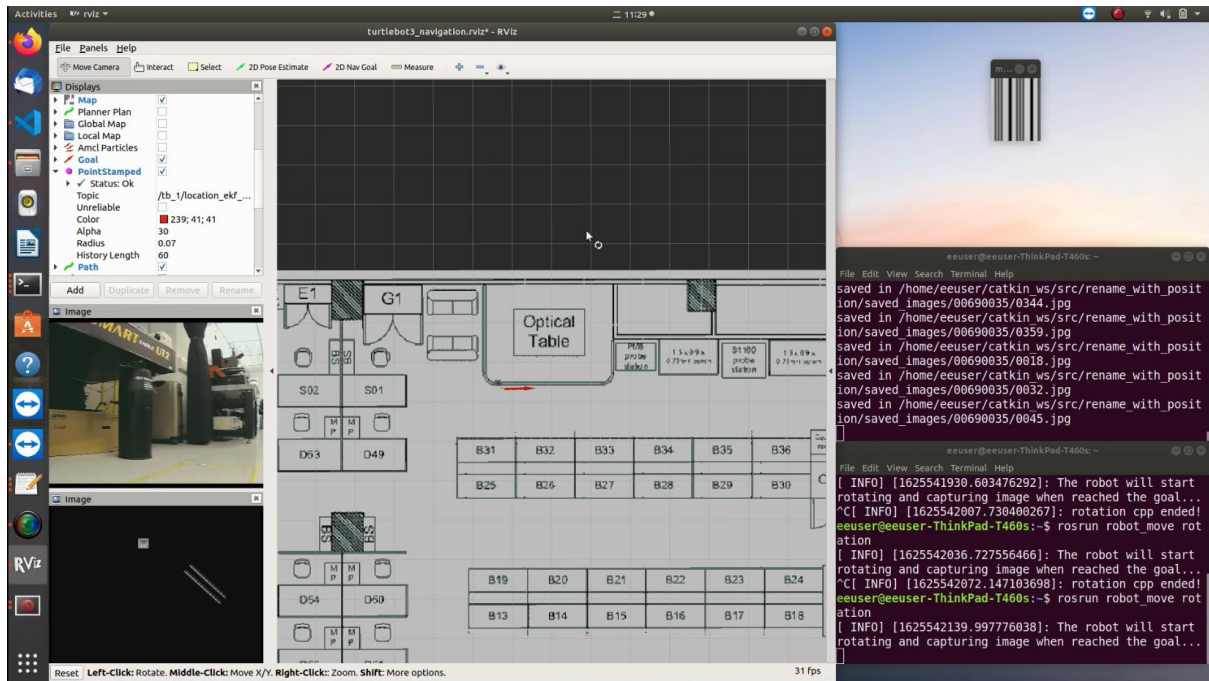
Figure 4.6: Diagram of the proposed VLP-based image taking robot system.

Figure 4.7 visualizes the node relationship in the proposed VLP-based image taking robot system using `rqt_graph` package in ROS system. The `/mkdir_position` node subscribes to three topics: the image view of the USB camera `/usb_cam/image_raw`, the status if the robot has reached the goal `/move_base/action_topics` and the location of the robot estimated by the proposed VLP system `/slovp_ekf_info`. Moreover, when the robot reaches the goal sent the by the controller, the `/mkdir_position` node will publish a topic named `/stop_rotation`. Then a node named `/rotation` will subscribe the `/stop_rotation` topic and publish a new velocity control topic given by zero linear velocity and a fixed angular velocity, which is subscribed by the robot

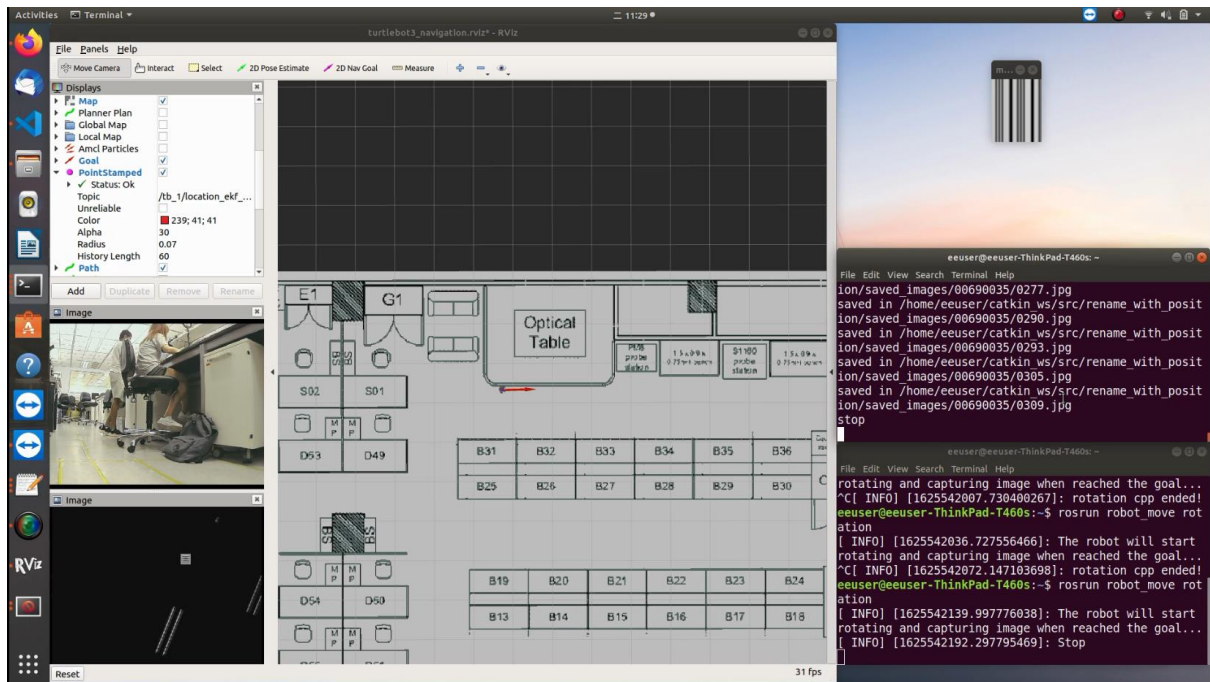
under the folder 01570039. When the robot has rotated for 360 degree, it will stop and wait for another command as shown in Figure 4.8(c). The terminal on the right bottom gives the stop instruction. Except for setting the navigation goal via RViz, we can send the coordinates of the target point using command lines. Figure 4.8(d) illustrates that a new goal whose coordinates are (0.06, 0.30, 0) in meters is sent to the robot via terminal.



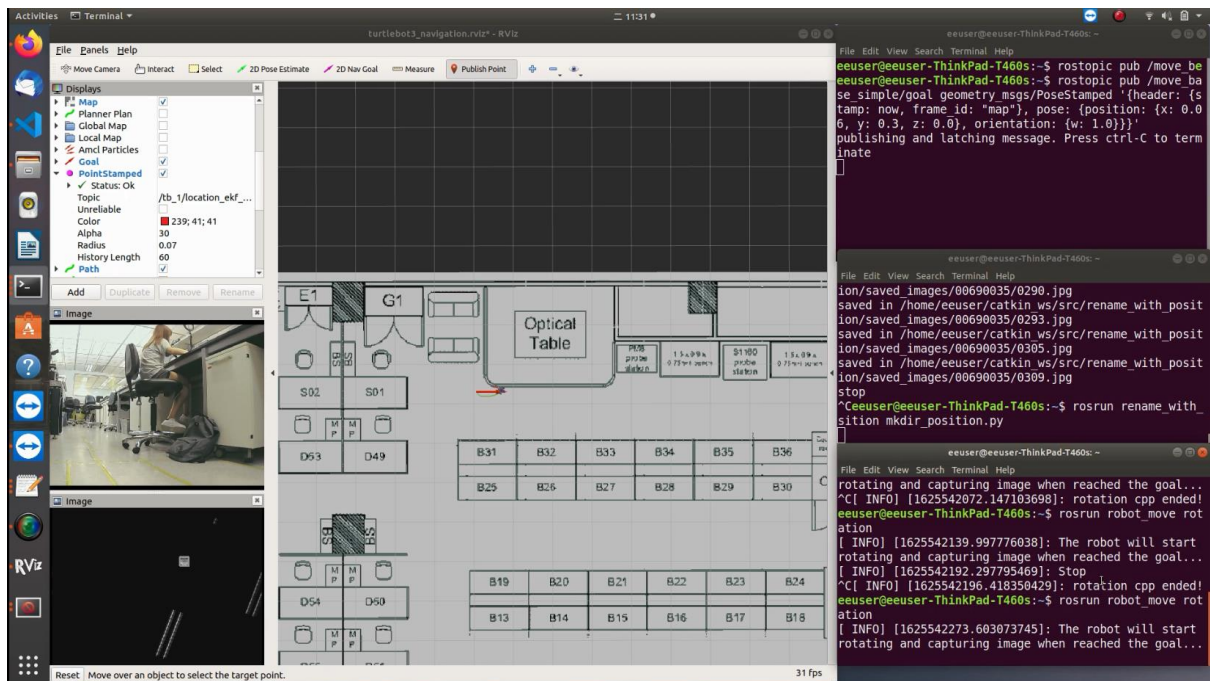
(a)



(b)



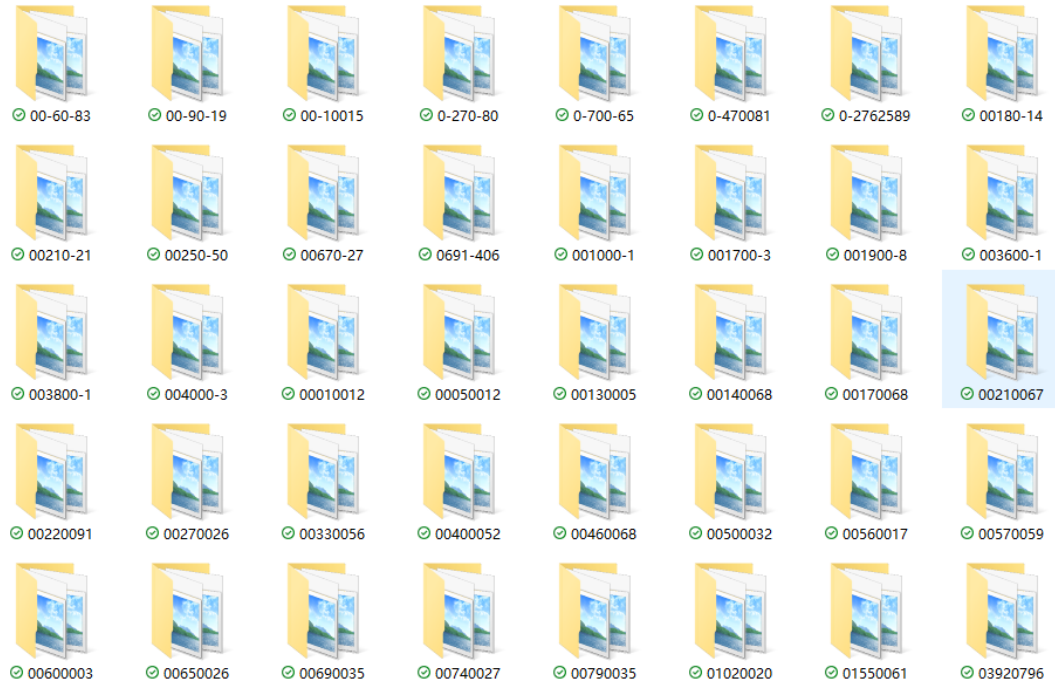
(c)



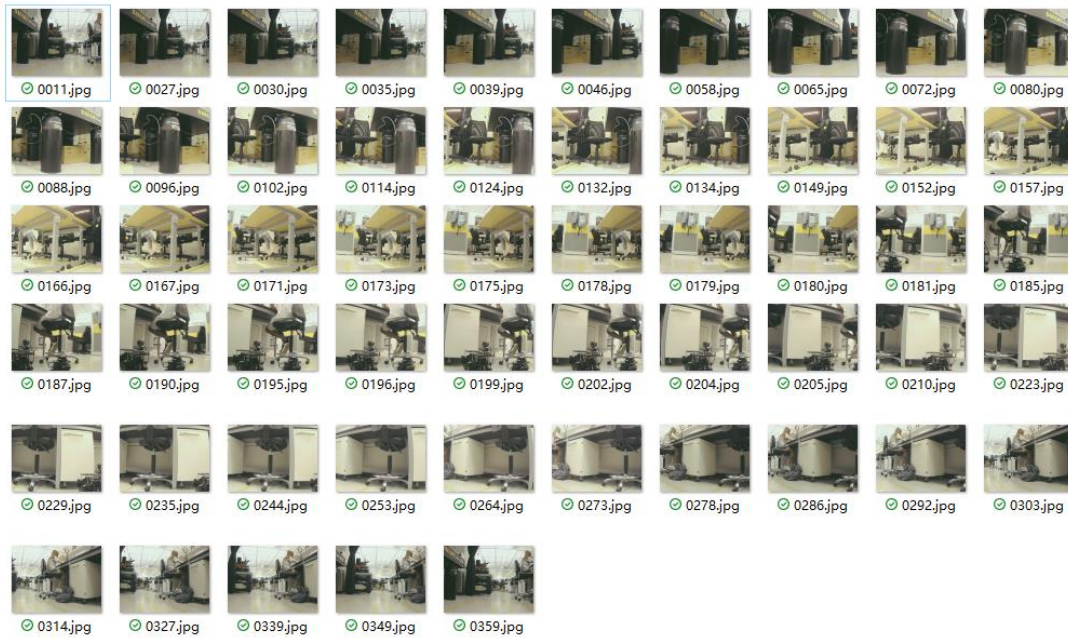
(d)

Figure 4.8: Navigation and image saving process: (a) sending a navigation goal to the robot via RViz, (b) the robot moving to the target point, (c) the robot saving images after arriving, (d) sending another navigation goal to the robot via a command line.

4.4.3 Panorama Creation



(a)



(b)

Figure 4.9: The saved images using the proposed system: (a) the created folders named with the x- and y-coordinates estimated by VLP, (b) the saved images named with the azimuth angle.

After the images are saved with the resolution of 640×480 in different folders for different locations as shown in Figure 4.9, we use opencv package to stitch the image. The stitching process follows the Brown and Lowe method [27], which is insensitive to the order of the images, the orientation of the images, any illumination changes or noisy images. The generated panorama shows in Figure 4.10.



Figure 4.10: The generated panorama using the saved images with location information.

4.5 Summary

This real-time robotic localization platform deals with high precision indoor positioning and VLC systems, both topics being hot trending in both academia and industry. It implements a completed positioning system with LEDs as transmitter and the camera on the robot as receiver and the position information is shown on the remote terminate in real time. The proposed system design will stimulate a wide range of the innovative utilization of VLP technology and provide a new idea for microwave or photonics-based positioning systems. Based on the proposed robot positioning and navigation system, a panorama creation method is proposed which can generate a panorama at any target point using a robot mounted with a typical USB camera.

4.6 References

- [1] W. Zhang, M. I. S. Chowdhury, and M. Kavehrad, "Asynchronous indoor positioning system based on visible light communications," *J. Opt. Eng.*, vol. 53, no. 4, pp. 1–9, 2014.
- [2] W. Guan, Y. Wu, S. Wen, H. Chen, C. Yang, Y. Chen and Z. Zhang, "A novel three-dimensional indoor positioning algorithm design based on visible light communication," *J. Opt. Commun.*, vol. 392, pp. 282-293.3, 2017.

- [3] E. I. Al Khatib, M. A. K. Jaradat, and M. F. Abdel-Hafez, "Lowcost reduced navigation system for mobile robot in indoor/outdoor environments," *IEEE Access*, vol. 8, pp. 25 014–25 026, 2020.
- [4] W. Lv, Y. Kang, and J. Qin, "Indoor localization for skid-steering mobile robot by fusing encoder, gyroscope, and magnetometer," *IEEE Trans. Syst. Man Cybern. Syst.*, vol. 49, no. 6, pp. 1241–1253, 2017.
- [5] Y. Murata and T. Murakami, "Estimation of posture and position based on geometric calculation using IMUs," in *Proc. IECON 2019-45th Annual Conference of the IEEE Industrial Electronics Society*, vol. 1. IEEE, 2019, pp. 5388–5393.
- [6] A. McGregor, G. Dobie, N. R. Pearson, C. N. MacLeod, and A. Gachagan, "Determining position and orientation of a 3-wheel robot on a pipe using an accelerometer," *IEEE Sens. J.*, vol. 20, no. 9, pp. 5061–5071, 2020.
- [7] F. A. Da Mota, M. X. Rocha, J. J. Rodrigues, V. H. C. De Albuquerque, and A. R. De Alexandria, "Localization and navigation for autonomous mobile robots using Petri nets in indoor environments," *IEEE Access*, vol. 6, pp. 31 665–31 676, 2018.
- [8] J. Zhang, Y. Lyu, J. Patton, S. C. Periaswamy, and T. Roppel, "BFVP: A probabilistic UHF RFID tag localization algorithm using Bayesian filter and a variable power RFID model," *IEEE Trans. Ind. Electron.*, vol. 65, no. 10, pp. 8250–8259, 2018.
- [9] D. Su, H. Kong, S. Sukkarieh, and S. Huang, "Necessary and sufficient conditions for observability of SLAM-based TDOA sensor array calibration and source localization," *IEEE Trans. Robot.*, 2021.
- [10] C. Naab and Z. Zheng, "Application of the unscented Kalman filter in position estimation a case study on a robot for precise positioning," *Rob. Auton. Syst.*, vol. 147, p. 103904, 2022.
- [11] G. Schouten and J. Steckel, "A biomimetic radar system for autonomous navigation," *IEEE Trans. Robot.*, vol. 35, no. 3, pp. 539–548, 2019.

- [12]S. Xu, W. Chou and H. Dong, "A robust indoor localization system integrating visual localization aided by CNN-based image retrieval with Monte Carlo localization," *Sensors*, vol. 19, (2), pp. 249, 2019.
- [13]C. Hsu, S. Liu, F. Lu, C. Chow, C. Yeh and G. Chang, "Accurate indoor visible light positioning system utilizing machine learning Technique with height tolerance," in *Proc. OFC 2018*, San Diego, CA, USA, 2018.
- [14]W. Guan, S. Chen, S. Wen, Z. Tan, H. Song and W. Hou, "High-accuracy robot indoor localization scheme based on robot operating system using visible light positioning," *IEEE Photon. J.*, vol. 12, no. 2, pp. 1-16, 2020.
- [15]E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numerische Math.*, vol. 1, no. 1, pp. 171-269, 1959.
- [16]D. -D. Zhu and J. -Q. Sun, "A New Algorithm Based on Dijkstra for Vehicle Path Planning Considering Intersection Attribute," *IEEE Access*, vol. 9, pp. 19761-19775, 2021.
- [17]P. E. Hart, N. J. Nilsson and B. Raphael, "A Formal Basis for the Heuristic Determination of Minimum Cost Paths," *IEEE Trans. Syst. Man Cybern. Syst.*, vol. 4, no. 2, pp. 100-107, 1968.
- [18]S. Quinlan and O. Khatib, "Elastic bands: connecting path planning and control," in *[1993] Proceedings IEEE International Conference on Robotics and Automation*, vol. 2, pp. 802-807, 1993.
- [19]C. Roesmann, W. Feiten, T. Woesch, F. Hoffmann and T. Bertram, "Trajectory modification considering dynamic constraints of autonomous robots," in *Proc. ROBOTIK 2012; 7th German Conference on Robotics*, pp. 1-6, 2012.
- [20]D. Fox, W. Burgard and S. Thrun, "The dynamic window approach to collision avoidance," *IEEE Robot. Autom. Mag.*, vol. 4, no. 1, pp. 23-33, March 1997.
- [21]W. Guan, L. Huang, B. Hussain and C. Patrick Yue, "Robust robotic localization using visible light positioning and inertial fusion," *IEEE Sens. J.*, doi: 10.1109/JSEN.2021.3053342.

- [22]B. Hussain, C. Qiu, and C. P. Yue, “A universal VLC modulator for retrofitting LED lighting and signage,” in *Proc. 2019 IEEE 8th Global Conference on Consumer Electronics (GCCE)*, Osaka, Japan, 2019, pp. 1008-1009.
- [23]C. Qiu, B. Hussain and C. P. Yue, “Bluetooth based wireless control for iBeacon and VLC enabled lighting,” in *Proc. 2019 IEEE 8th Global Conference on Consumer Electronics (GCCE)*, Osaka, Japan, 2019, pp. 614-615.
- [24]B. Hussain, C. Lau, and C. P. Yue, “Li-Fi based secure programmable QR code (LiQR),” in *JSAP-OSA Joint Symposia*. Fukuoka, Japan: Optical Society of America, 2017, p. 6p_A409_6.
- [25]W. Guan, X. Zhang, Y. Wu, Z. Xie, J. Li and J. Zheng, “High precision indoor visible light positioning algorithm based on double LEDs using CMOS image sensor,” *Appl. Sci.*, vol. 9, no. 6, pp. 1238, 2019.
- [26]G. Schouten and J. Steckel, “A biomimetic radar system for autonomous navigation,” *IEEE Trans. Robot.*, vol. 35, no. 3, pp. 539–548, 2019.
- [27]M. Brown, D. G. Lowe, “Automatic panoramic image stitching using invariant features,” *Int. J. Comput. Vis.*, vol. 74, no. 1, pp. 59-73, 2007.

CHAPTER 5 VLP and SLAM-assisted Map Calibration for Robot Navigation

5.1 Introduction

With the development of sensors, control systems, bionics and artificial intelligence, robot technology has been investigated and applied in many areas to provide services such as hospital inspection, hotel delivery and warehouse logistics. Using mobile robots in indoor environments can effectively improve the intelligence and effectiveness of task execution. Meanwhile, in robot applications, navigation plays an increasingly crucial role. As an essential element in the navigation process, high-precision positioning in indoor environments is still a challenging task. Since the Global Positioning System (GPS) can not provide satisfactory positioning services in indoor environments due to the extreme attenuation and interruption caused by indoor structures, Wi-Fi/Bluetooth fingerprinting-based indoor positioning systems (IPs) have raised extensive attention and achieved encouraging results. However, positioning based on Wi-Fi/Bluetooth can only achieve meter-level accuracy [1].

Compared with Wi-Fi/Bluetooth fingerprinting-based positioning, positioning with landmarks composed of visible light positioning (VLP)-enabled LED lights can provide an absolute location when using an image sensor as a receiver. Scanning of the whole area is not required, and global 3D positioning results can be achieved as long as the 3D position information of the landmarks is encoded in the VLP lights. In our previous works [2] and [3], we proposed a VLP system based on a single LED that could achieve centimeter-level accuracy, with an average accuracy of 2.1 cm for a stationary robot [2] and of 3.9 cm for a 3D tilted receiver camera [3].

Besides positioning, building an accurate map is another important element for navigation because both positioning and path planning rely on the map information of the environment [4]. One typical map representation is an occupancy grid map [5], in which the value of each cell represents the probability of being occupied by obstacles. Currently, Simultaneous Localization and Mapping (SLAM) technology [6] is widely used to determine the position of

the robot and build an occupancy grid map at the same time by fusing available sensor information. However, SLAM has three main drawbacks. The first is that SLAM can only determine a local position and the relative movement of the robot in the environment. Second, SLAM requires scanning and survey of the whole scene to get a map. Last, the position based on the sensors in the robot, including the odometer and inertial measurement unit (IMU) will drift and lose global accuracy with time. These drawbacks lead to challenges for mapping and navigation in large-scale and multi-floor environments.

In addition, the occupancy grid map created by SLAM only contains three types of information: the cell is occupied, free of obstacles or unknown to the robot. There is no semantic information of the structures. Moreover, the occupancy grid map may not be oriented so that humans can distinguish the direction with a correspondence to the real world. The noise from the sensors will also be shown on the map and mislead the robot as well as humans. Therefore, it is noticeably difficult for humans to understand an occupancy grid map generated by a robot and send commands to the robot based on it.

Noting the drawbacks of the occupancy grid map generated by SLAM and its difficulties for humans, we propose to use a layout map to promote better cooperation between humans and robots. A human in an indoor environment will always use a layout map, which illustrates structures and contains semantic information, to navigate a pathway from the current position to the target position. A layout map always demonstrates the whole area and is complete and without noise. The boundaries on such maps refer to obstacles in the area that can not be crossed by a robot and have the same meaning as the occupied cells in an occupancy grid map. However, the accuracy of a layout map in terms of resolution can not be guaranteed, which will degrade the accuracy and reliability of navigation.

Therefore, in this chapter, we propose to calibrate the layout map of a scene using the occupancy grid map generated by SLAM to improve navigation performance. In the mapping process, an image sensor is mounted on the robot and we use VLP landmarks to acquire the robot's position on the layout map. At the same time, SLAM is performed on the robot, and its position on the occupancy grid map is determined by the sensors. After at least two landmarks

are tracked by the image sensor on the robot, the occupancy grid map generated by SLAM is saved as a sensor map. Then, the orientations of the two maps are aligned based on the pixel coordinates of the tracked landmarks on the maps. Moreover, the scale of the layout map is calibrated by computing the pixel distance between the key points. To keep the consistency of the map image after scaling, the eight-neighborhood averaging method and bilinear interpolation method are applied. It is noteworthy that the map calibration method based on landmarks is scalable to scenes mounted with multiple landmarks by computing the average of the rotation angle and scale of every two key points. We finally present experiments to verify the effectiveness of the proposed map alignment methods and analyze the performance on the robot operating system (ROS).

The chapter is organized as follows. Related work is introduced in Section 5.2. Section 5.3 explains the mapping system and map transformation process. In Section 5.4, we present the details of the proposed VLP landmarks and SLAM-assisted automatic map calibration method. Experimental results are provided and analyzed in Section 5.5. Finally, Section 5.6 concludes this chapter.

5.2 Related Work

To create a complete map of an environment, especially a large venue, a map merging method is widely used to integrate the occupancy grid maps generated at different locations of the environment to be perceived. A map merging method based on pose graphs is presented in [7], which requires consecutive pose information of the robot to remove the distortion of the generated maps. However, due to the accumulated error from robot sensors, it is difficult to continuously obtain high-accuracy positioning results without landmark-based error correction, especially in a large venue. In [8], a pair-wise map merging method is proposed to integrate the local maps built by different robots into a single global map. However, it requires high overlapping percentage between two maps, otherwise it will lead to unreliable map integration performance. Multi-robot cooperative mapping by introducing augmented variables to parallelize the computation is proposed in [9], while in [10], a robust map merging algorithm with multi-robot SLAM (MRSLAM) is proposed, but it also requires a large amount of overlap

between the two maps generated by two robots to extract and match the features in the two maps. In [11], the existing methods on merging redundant line segments are evaluated by experiments. A distributed method for constructing an occupancy grid map using a swarm of robots with global localization capabilities and limited inter-robot communication is proposed in [12] and physical experiments are performed. Instead of a diffusive random walk of the robots, Lévy walks and larger individual memory are applied to the robots. The drawback of all these map merging methods, however, based on a single robot or multiple robots, is that they rely on scanning the whole environment to get complete map information, which is time-consuming and certainly leads to a high cost.

During recent years, a significant amount of work has been done on map alignment of different types of maps. In [13], a map alignment method for a floor map and an occupancy grid map generated by SLAM using a similarity transformation is proposed. The process is not time-consuming, but it has poor performance on maps with noise, different scales or types of maps. An improved SLAM using the Bayesian prior extracted from a blueprint is presented in [14]. It improves the performance of SLAM algorithm, but in order to determine the correspondence of two kinds of maps, the semantic information on the layout map has to be eliminated. Therefore, it is still difficult for humans to understand the generated map, and the method cannot actually facilitate the collaboration between humans and robots. Scanning of the whole scene using SLAM is also required. A nonlinear optimization method for nonrigid alignment of maps is proposed in [15], but it has a high computation cost due to the nonlinear optimization. A fast map matching algorithm based on area segmentation is presented in [16]. However, it also requires scanning the whole area and is sensitive to the occupancy grid maps with distortion induced by the accumulated error from robot sensors.

Therefore, calibrating and aligning maps of different types and maps with distortions or noise is still a challenging task. In this chapter, we propose to calibrate a layout map with a sensor map generated by SLAM. The proposed method works for maps in different orientations or scales. We use high-accuracy VLP landmarks to obtain the position of the robot on the layout map and align it with its position on the sensor map. At least two landmarks are placed in the

environment to be perceived, and therefore scanning the whole venue is not required. With a calibrated layout map, a human can send instructions to the robot with the semantic information shown on the map and the robot can navigate to the target point with the aid of the occupancy information on the map, by which the efficiency thus is improved.

5.3 Mapping System

In this section, the system design of the proposed map calibration technology, including the occupancy grid mapping system, map transformation method and the proposed diagram, will be presented.

5.3.1 Occupancy Grid Mapping System

An occupancy grid map, which consists of an array of cells representing the occupancy information of an environment, was first introduced in [17] and is usually generated from SLAM. The binary variable in each cell represents the probability of the presence of an obstacle at that location of the perceived environment. If the variable is closer to 0, there is a higher certainty that the cell is not occupied and is free of obstacles. If the variable is 0.5, the cell is unknown to the robot, neither occupied nor free. The probability in each cell is relatively independent. An occupancy map is updated by the detection results from robot sensors. In this chapter, we use a 2D occupancy map to describe a slice of the 3D perceived environment. When we save the occupancy grid map as an image file, the probability in each cell p will convert to a grayscale value in each pixel g :

$$g = -254p + 254. \quad (5-1)$$

Therefore, if the probability of an obstacle in the cell is close to 0, the grayscale value in that pixel will be close to 254, indicated in white. Otherwise, the color of the pixel will reach black.

5.3.2 Map Transformation

To achieve map calibration, map transformation, including rotation, translation and scaling, will be performed on the original map image. For a grayscale map image $\mathbf{G}$ of the size $h \times w$, each pixel contains the grayscale value of that pixel. As the grayscale values represent three

different occupancy meanings to the robot, we divide the values in matrix $\mathbf{G}$ to three units, occupied, unknown and free of obstacles, with two thresholds given by t_o and t_f , as shown in Figure 5.1. Thus, if the grayscale value is lower than t_o , the pixel is occupied. If the grayscale value of the pixel is higher than t_f , the pixel is free of obstacles. Otherwise, the pixel is unknown for the robot.

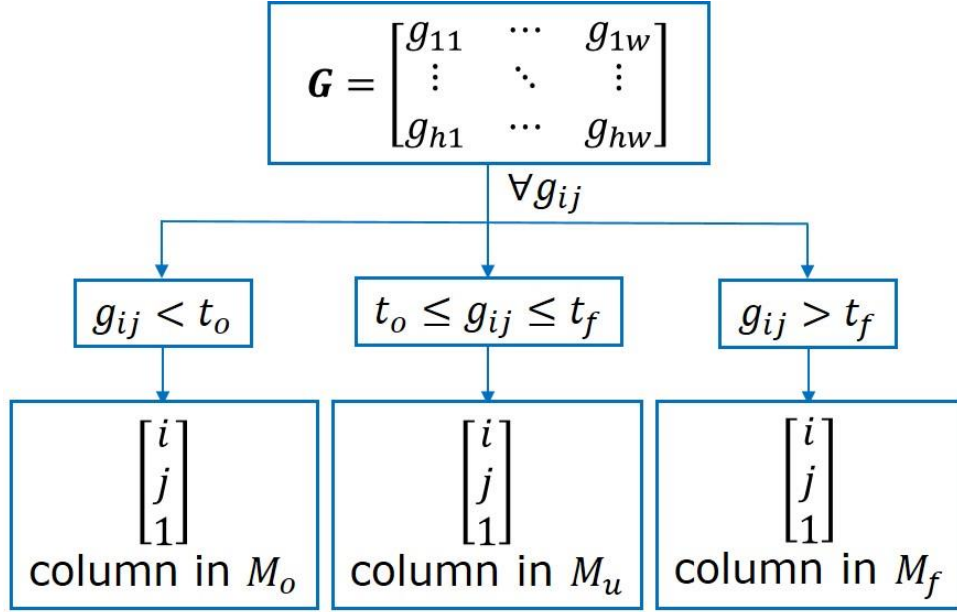


Figure 5.1: Occupancy-coordinate transformation for a grayscale map image.

To present the map transformation method in an intuitive way, we represent the map image as three matrices, $\mathbf{M}_o$, $\mathbf{M}_u$ and $\mathbf{M}_f$, containing the pixel coordinates of the grayscale values in the three units mentioned above. The three matrices are of the size $3 \times N_o$, $3 \times N_u$ and $3 \times N_f$, respectively, where N_o is the number of pixels that are occupied, N_u is the number of pixels that are unknown to the robot, and N_f is the number of pixels that are free of obstacles, and $N_o + N_u + N_f = hw$. The first rows in the three matrices represent the u-coordinate of the pixels, and the second rows represent the v-coordinate of the pixels in the pixel coordinate system. All the values in the last rows are assigned '1'. The order in which we place the pixel coordinates of the grayscale values in matrix $\mathbf{G}$ in the occupancy-coordinate matrices is based on checking the grayscale values in $\mathbf{G}$ row by row and then placing their pixel coordinates into the corresponding matrix.

Map transformation can be directly performed by matrix multiplication on the occupancy-coordinate matrices $\mathbf{M}_o$, $\mathbf{M}_u$ and $\mathbf{M}_f$. For example, if the map is supposed to rotate clockwise with angle φ at the center (w_c, h_c) , enlarge by k times, and then translate (w_t, h_t) , we should firstly translate the origin of the coordinate system to the map center (w_c, h_c) , and the first translation matrix $\mathbf{T}_c$ is given by

$$\mathbf{T}_c = \begin{bmatrix} 1 & 0 & -w_c \\ 0 & 1 & -h_c \\ 0 & 0 & 1 \end{bmatrix}. \quad (5-2)$$

Subsequently, we should rotate the map clockwise with angle φ , and the rotation matrix $\mathbf{R}$ is given by

$$\mathbf{R} = \begin{bmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (5-3)$$

The next step is to enlarge the image, where the scaling matrix $\mathbf{S}$ can be described as

$$\mathbf{S} = \begin{bmatrix} k & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (5-4)$$

Finally, we will translate the origin of the coordinate system back and further translate (w_t, h_t) , and the translation matrix $\mathbf{T}$ is given by

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & k * w_c + w_t \\ 0 & 1 & k * h_c + h_t \\ 0 & 0 & 1 \end{bmatrix}. \quad (5-5)$$

Therefore, the transformed occupancy-coordinate matrices $\mathbf{M}_{o_t}$, $\mathbf{M}_{u_t}$ and $\mathbf{M}_{f_t}$ can be obtained by

$$\begin{aligned} \mathbf{M}_{o_t} &= \mathbf{TSRT}_c \mathbf{M}_o, \\ \mathbf{M}_{u_t} &= \mathbf{TSRT}_c \mathbf{M}_u, \\ \mathbf{M}_{f_t} &= \mathbf{TSRT}_c \mathbf{M}_f. \end{aligned} \quad (5-6)$$

It is noteworthy that, after we scale up an image, each pixel of the original map image is moved in a certain direction based on the scaling constant k . However, if the scaling factor is larger than 1, there may exist unassigned pixel values in the resultant map image, which are regarded as holes. Furthermore, if the scaling factor is smaller than 1, there will be multiple assigned pixels. Therefore, we will add an interpolation and eight-neighbourhood averaging

method after scaling transformation to appropriately assign the grayscale values to these pixels. The details will be described in Section 5.4.3.

5.3.3 Overview of the Proposed Map Calibration System

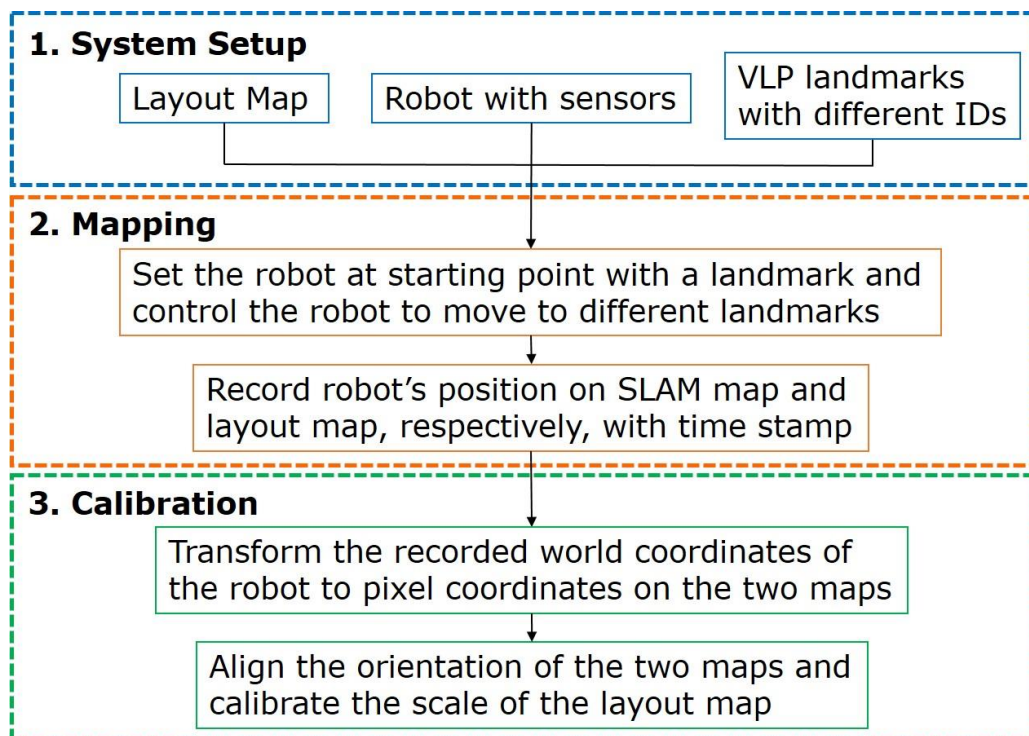


Figure 5.2: The proposed VLP landmarks and SLAM-assisted automatic map calibration for robotic navigation: (1) system setup, (2) mapping process, (3) calibration process.

The diagram of the proposed map calibration system using VLP landmarks and SLAM is given in Figure 5.2. The system setup contains a layout map of the environment to be perceived, a robot, and multiple VLP lights with different IDs installed in the experimental area. The VLP landmarks are mounted on the ceiling and controlled by visible light communication (VLC)-enabled light emitting diode (LED) drivers to transmit optical signals [18]. The LEDs are modulated by the on-off keying scheme and encoded with unique IDs, which contain the LEDs' world coordinates stored in a uniform resource identifier (URI) database. The robot used in the mapping system is equipped with multiple sensors, including an IMU, odometer and LiDAR. A camera is also mounted on the robot to face toward the ceiling and works in rolling shutter mode to capture the signals broadcasted by the VLP-based LEDs, decode and extract the position information.

To start the mapping process, the robot is set under a VLP landmark and uses the camera to capture the LED image and decode the position of the starting point. Then we apply SLAM to the robot and control it to move to different landmarks and record its positions, as acquired from SLAM and VLP landmarks, respectively. The time stamp is simultaneously marked with the position. After the robot has tracked at least two landmarks in the area, the mapping process can be stopped and the occupancy grid map generated by SLAM is saved. Therefore, we obtain one sensor map with the robot's positions from SLAM and one layout map with the robot's positions from VLP landmarks. Subsequently, we propose to use the obtained sensor map and the robot's positions on two maps to calibrate the layout map. Firstly, we transform the recorded world coordinates of the robot to pixel coordinates on the two different maps. Then we calibrate the layout map by aligning the robot's positions on it and the occupancy map.

5.4 Map Calibration Method

In this section, we will describe the details of the proposed VLP landmarks and SLAM-assisted map calibration method. As we mentioned, a layout map contains semantic information, which is readable for humans to give instruction to robots. However, the scale of a layout map may not be accurate, leading to an inaccurate resolution of the map in terms of meters per pixel. In a large scene which is to be perceived, it is difficult and complex to get the resolution through measurement. Compared with a layout map, the sensor map generated from SLAM has a much more accurate resolution, but more noise points. Therefore, we propose to calibrate the scale of the layout map, which will help robots to achieve better navigation performance.

5.4.1 Positioning on Two Different Maps

Figure 5.3 illustrates the proposed mapping process. At least two VLP landmarks are required to be mounted on the ceiling of the environment which is to be perceived. The exact position of the mounting location in the environment is encoded in the VLP light and is broadcasted to the robot by OOK modulation. The positions of the VLP lights are also marked on the layout map. We use a robot equipped with a camera, odometer, IMU and LiDAR. The camera is set to face the ceiling and is used to get the robot's position by VLP. When the camera

detects a VLP landmark, it will decode the position information encoded in the rolling shutter patterns, and then translate it to its own position. The translation from the world coordinates of the landmark to the world coordinates of the robot is calculated by the location of the landmark on the camera plane and the orientation of the robot determined by the odometer on the robot. Then the world coordinates of the robot are further translated to the pixel coordinates on the layout map as the landmarks are labeled on the layout map.

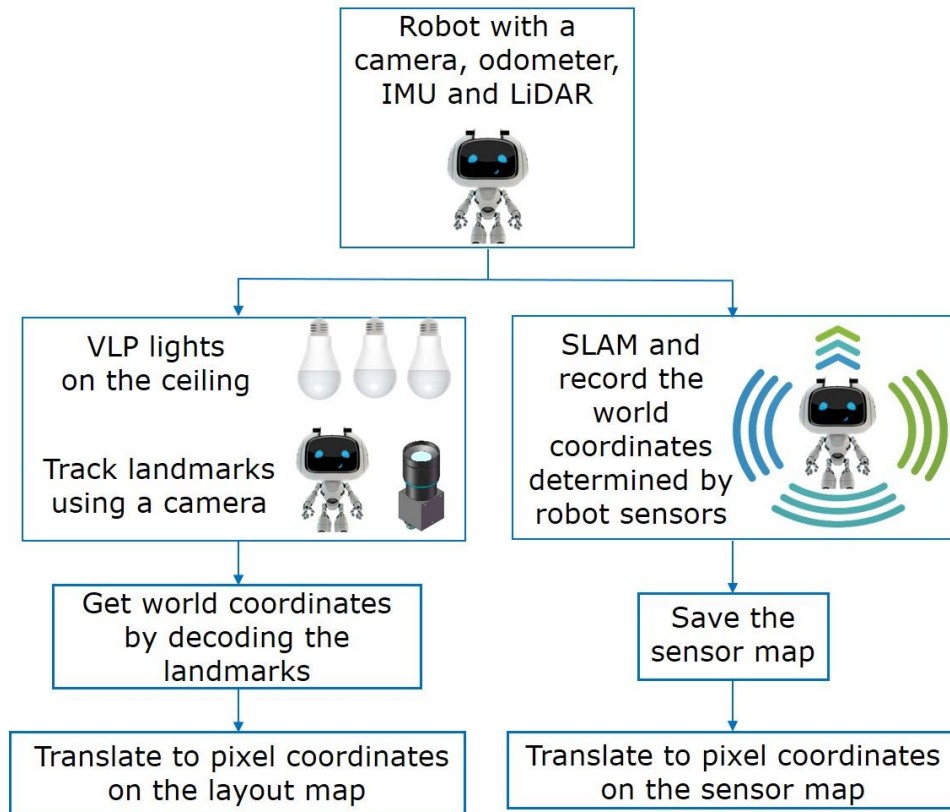


Figure 5.3: Mapping process with VLP landmarks and SLAM technology.

In the mapping process, the robot starts under one of the landmarks to get the first key point for map calibration. Then it is controlled to perceive the environment and conduct SLAM with its odometer, IMU and LiDAR sensors. The world coordinates obtained from the robot sensors are recorded at the same time. After the robot has tracked at least two VLP landmarks, which means that we get at least two pairs of coordinates of the key points from the two different positioning methods, we can save the occupancy grid map created by SLAM as a sensor map, set the resolution of the map in terms of meters per pixel, and then translate the recorded world coordinates of the robot to pixel coordinates on the sensor map with the resolution.

5.4.2 Calibration of the Orientation

A layout map is readable for humans and is always presented in an orientation in which humans can understand the semantic information. However, the sensor map created by the robot may not always be in the same orientation as the layout map used by a human. Therefore, we propose to correct the orientation of the sensor map.

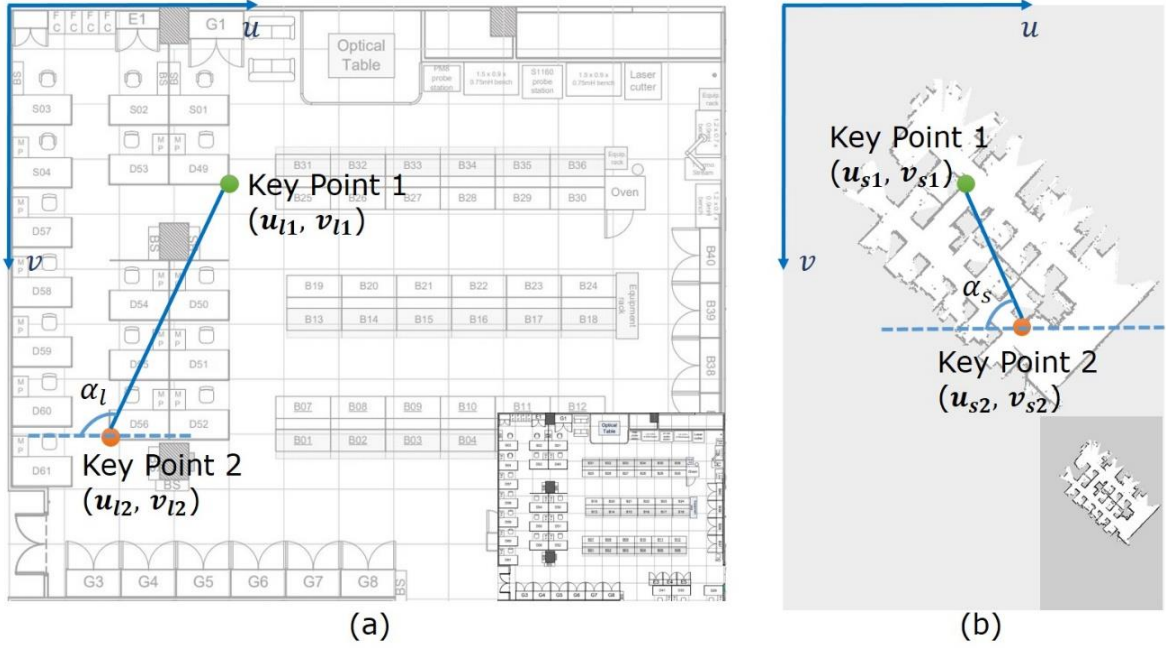


Figure 5.4: Key points on the two maps: (a) layout map, (b) sensor map.

Firstly, we convert the layout map to a grayscale image $\mathbf{G}_l$. Since in a typical layout map, furniture and structures are drawn with black or dark squares. Thus, the grayscale values in the converted layout map have the same meaning as the grayscale values in the sensor map, where if the pixel is in black and its grayscale value is close to 0, the probability of an obstacle at that point is close to 1. Then for the saved sensor map $\mathbf{G}_s$, we firstly find the occupancy-coordinate representation given by $\mathbf{M}_{s_o}$, $\mathbf{M}_{s_u}$ and $\mathbf{M}_{s_f}$. Then in the mapping process, we assume that the robot has detected two landmarks and labeled them on the two maps according to the translated pixel coordinates given by (u_{l1}, v_{l1}) and (u_{l2}, v_{l2}) on the layout map and (u_{s1}, v_{s1}) and (u_{s2}, v_{s2}) on the sensor map, as shown in Figure 5.4. Then we draw a line between the two key points and find the angle between the line and the negative u-axis in the pixel coordinate system.

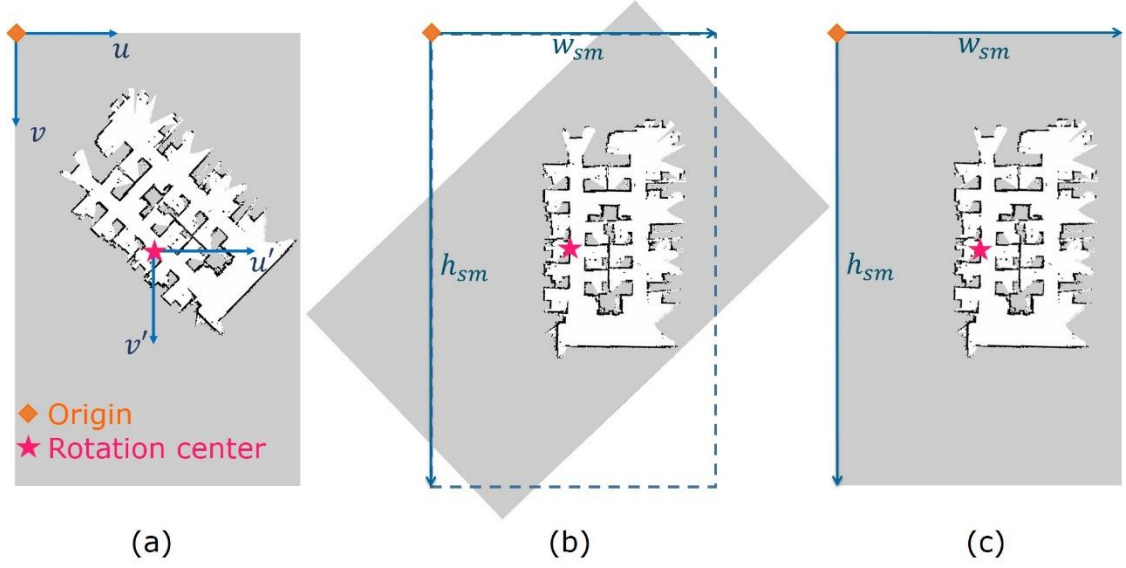


Figure 5.5: Orientation calibration for the sensor map: (a) translate the rotation center, (b) rotate and crop, (c) fill in the corners.

To rotate the sensor map clockwise with the angle of $\alpha_l - \alpha_s$, where α_l is the angle on the layout map and α_s is the angle on the sensor map, we firstly translate the origin of the coordinate system to the map center $(\frac{w_s}{2}, \frac{h_s}{2})$, which is the rotation center, as shown in Figure 5.5(a). The origin translation matrix T_{s1} is given by

$$T_{s1} = \begin{bmatrix} 1 & 0 & -\frac{w_s}{2} \\ 0 & 1 & -\frac{h_s}{2} \\ 0 & 0 & 1 \end{bmatrix}, \quad (5-7)$$

where w_s and h_s are the width and height of the sensor map image G_s . Subsequently, we substitute the rotation angle $\alpha_l - \alpha_s$ into the rotation matrix defined in (5-3) as

$$R_s = \begin{bmatrix} \cos(\alpha_l - \alpha_s) & -\sin(\alpha_l - \alpha_s) & 0 \\ \sin(\alpha_l - \alpha_s) & \cos(\alpha_l - \alpha_s) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (5-8)$$

Then we translate the origin of the coordinate system back, and the translation matrix T_{s2} is given by

$$T_{s2} = \begin{bmatrix} 1 & 0 & \frac{w_s}{2} \\ 0 & 1 & \frac{h_s}{2} \\ 0 & 0 & 1 \end{bmatrix}. \quad (5-9)$$

By multiplying the translation matrices and rotation matrix, the occupancy-coordinate representation matrices of the sensor map are given by

$$\begin{aligned} \mathbf{M}_{s_{ot}} &= \mathbf{T}_{s2} \mathbf{R}_s \mathbf{T}_{s1} \mathbf{M}_{s_o}, \\ \mathbf{M}_{s_{ut}} &= \mathbf{T}_{s2} \mathbf{R}_s \mathbf{T}_{s1} \mathbf{M}_{s_u}, \\ \mathbf{M}_{s_{ft}} &= \mathbf{T}_{s2} \mathbf{R}_s \mathbf{T}_{s1} \mathbf{M}_{s_f}, \end{aligned} \quad (5-10)$$

where $\mathbf{M}_{s_{ot}}$, $\mathbf{M}_{s_{ut}}$ and $\mathbf{M}_{s_{ft}}$ are the occupancy-coordinate matrices after rotation. Then we assign the corresponding grayscale values to the pixels in sensor map image $\mathbf{G}_s$ and get the rotated sensor map image $\mathbf{G}_{st}$.

After rotation, the rotated sensor map image $\mathbf{G}_{st}$ may not be in the original size of the sensor map image given by $w_s \times h_s$. Therefore, we find the maximal u-coordinate of the pixels indicating the cells are free of obstacles or occupied by an obstacle, which is given by

$$\begin{aligned} p_{o_{um}} &= \max[\mathbf{M}_{s_{ot}}(1, i)], i \in [1, N_o] \cap \mathbb{Z}, \\ p_{f_{um}} &= \max[\mathbf{M}_{s_{ft}}(1, j)], j \in [1, N_f] \cap \mathbb{Z}, \end{aligned} \quad (5-11)$$

where $\mathbb{Z}$ is the integer set, N_o is the width of matrix $\mathbf{M}_{s_{ot}}$ and N_f is the width of matrix $\mathbf{M}_{s_{ft}}$. Similarly, we determine the maximal v-coordinate of the pixels indicating the cells are free of obstacles or occupied by an obstacle, which is given by

$$\begin{aligned} p_{o_{vm}} &= \max[\mathbf{M}_{s_{ot}}(1, i)], i \in [1, N_o] \cap \mathbb{Z}, \\ p_{f_{vm}} &= \max[\mathbf{M}_{s_{ft}}(1, j)], j \in [1, N_f] \cap \mathbb{Z}, \end{aligned} \quad (5-12)$$

Then we find the maximum in the u-coordinates and v-coordinates, respectively, given by

$$\begin{aligned} w_{sm} &= \max[w_s, p_{o_{um}}, p_{f_{um}}], \\ h_{sm} &= \max[h_s, p_{o_{vm}}, p_{f_{vm}}], \end{aligned} \quad (5-13)$$

where (w_{sm}, h_{sm}) is the size to which we will crop the rotated sensor map.

Then we trim the map image by the width of w_{sm} and the height of h_{sm} , as shown in Figure 5.5(b) and delete those columns in $\mathbf{M}_{s_{ot}}$, $\mathbf{M}_{s_{ut}}$ and $\mathbf{M}_{s_{ft}}$. Furthermore, we fill in the corners with the grayvalues indicating that the pixel is unknown, namely, the half probability of an obstacle, as shown in Figure 5.5(c), and add the pixel coordinates of the elements in the corners to matrix $\mathbf{M}_{s_{ut}}$. The obtained occupancy-coordinate matrices of the sensor map after cropping and filling are given by $\mathbf{M}_{s_{om}}$, $\mathbf{M}_{s_{um}}$ and $\mathbf{M}_{s_{fm}}$.

It is noteworthy that the map rotation process we describe above is based on two key points, namely, two VLP landmarks, but it is scalable to a perceived environment that is mounted with multiple landmarks by computing the average rotation angle of every two key points on the map and substituting the average angle into (5-8).

5.4.3 Calibration of the Scale

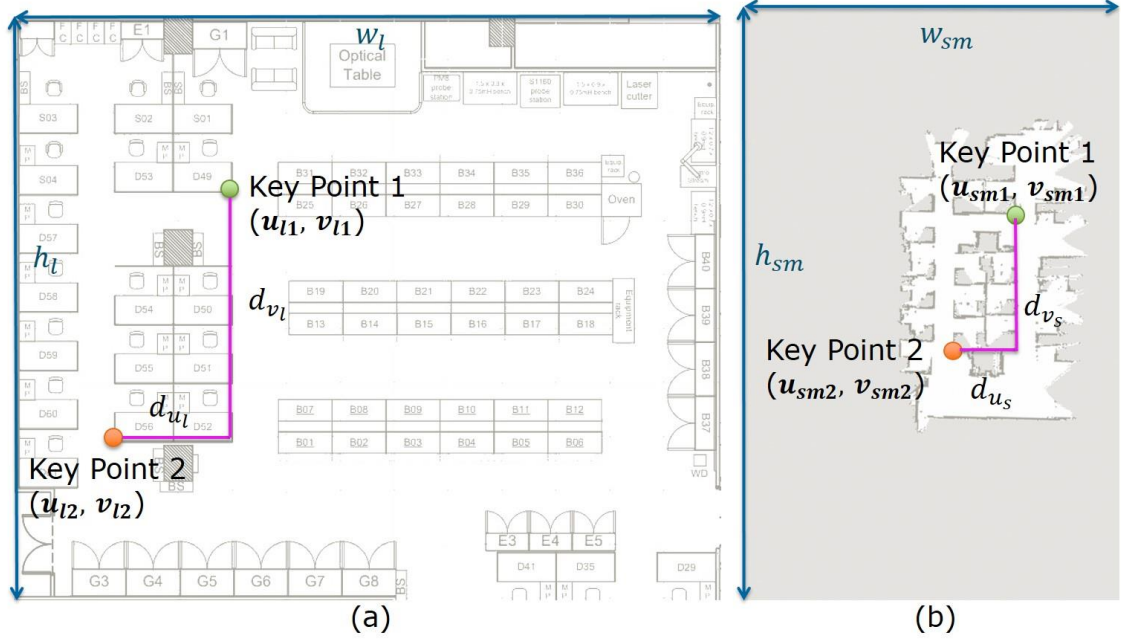


Figure 5.6: Vertical and horizontal distances between the two key points: (a) on the layout map, (b) on the sensor map.

After we align the orientation of the two maps, the next step is to align their scales. As the resolution of the sensor map obtained from SLAM is more accurate than that of a manually drawn map, we propose to calibrate the scale of the layout map with the sensor map. Firstly, we determine the occupancy-coordinate representation of the layout map matrix given by $\mathbf{M}_{l_o}$, $\mathbf{M}_{l_u}$ and $\mathbf{M}_{l_f}$, which indicate the pixels are occupied, unknown or free of obstacles, respectively. Then we find the pixel distance between the two key points on the layout map, given by d_{u_l} in the u-axis and d_{v_l} in the v-axis, as shown in Figure 5.6(a), where w_l is the width and h_l is the height of the layout map. Similarly, we get the pixel distance between the two key points on the sensor map, given by d_{u_s} in the u-axis and d_{v_s} in the v-axis, as shown

in Figure 5.6(b). It is noteworthy that, according to (5-10), the pixel coordinates of the two key points on the sensor map have also been transformed after the rotation as

$$\begin{aligned} \begin{bmatrix} u_{sm1} \\ v_{sm1} \\ 1 \end{bmatrix} &= \mathbf{T}_{s2} \mathbf{R}_s \mathbf{T}_{s1} \begin{bmatrix} u_{s1} \\ v_{s1} \\ 1 \end{bmatrix}, \\ \begin{bmatrix} u_{sm2} \\ v_{sm2} \\ 1 \end{bmatrix} &= \mathbf{T}_{s2} \mathbf{R}_s \mathbf{T}_{s1} \begin{bmatrix} u_{s2} \\ v_{s2} \\ 1 \end{bmatrix}, \end{aligned} \quad (5-14)$$

where (u_{sm1}, v_{sm1}) and (u_{sm2}, v_{sm2}) are the pixel coordinates of the key points on the sensor map after rotation.

Next, we modify the scale of the layout map, and the scaling matrix $\mathbf{S}_l$ is given by

$$\mathbf{S}_l = \begin{bmatrix} \frac{d_{u_l}}{d_{u_s}} & 0 & 0 \\ 0 & \frac{d_{v_l}}{d_{v_s}} & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (5-15)$$

Then, we multiply matrix $\mathbf{S}_l$ with the occupancy-coordinate representation matrices of the layout map as

$$\begin{aligned} \mathbf{M}_{l_{or}} &= \mathbf{S}_l \mathbf{M}_{l_o}, \\ \mathbf{M}_{l_{ur}} &= \mathbf{S}_l \mathbf{M}_{l_u}, \\ \mathbf{M}_{l_{fr}} &= \mathbf{S}_l \mathbf{M}_{l_f}, \end{aligned} \quad (5-16)$$

where $\mathbf{M}_{l_{or}}$, $\mathbf{M}_{l_{ur}}$ and $\mathbf{M}_{l_{fr}}$ are the occupancy-coordinate layout map matrices after scaling.

Since the scaling factor $\frac{d_{u_l}}{d_{u_s}}$ may not be an integer, the obtained pixel coordinates in the layout map matrices $\mathbf{M}_{l_{or}}$, $\mathbf{M}_{l_{ur}}$ and $\mathbf{M}_{l_{fr}}$ may result in non-integers. Therefore, we firstly round all the values in $\mathbf{M}_{l_{or}}$, as given by

$$\mathbf{M}_{l_{or}}(i, j) = \lfloor \mathbf{M}_{l_{or}}(i, j) + \frac{1}{2} \rfloor, \forall \mathbf{M}_{l_{or}}(i, j) \notin \mathbb{N}, \quad (5-17)$$

where $\lfloor \cdot \rfloor$ is the floor function, $i = 1, 2$, $j \in [1, N_{l_o}] \cap \mathbb{Z}$, N_{l_o} is the width of matrix $\mathbf{M}_{l_{or}}$ and $\mathbb{N}$ denotes the natural number set. Similarly, we round the values in $\mathbf{M}_{l_{ur}}$ and $\mathbf{M}_{l_{fr}}$.

In addition, if the scaling factor is smaller than 1, after multiplying the scaling matrix, there will be multiple columns in $\mathbf{M}_{l_{or}}$ composed of the same pixel coordinates. Then, for each occupancy-coordinate matrix, we treat each column as a single entity and extract the unique columns with no repetitions. The extracted columns constitute three new matrices given by

$M_{l_{oe}}$, $M_{l_{ue}}$ and $M_{l_{fe}}$. Furthermore, one pair of pixel coordinates may occur in different occupancy-coordinate matrices, indicating different grayscale values in the layout map image. To keep the consistency of the grayscale values in the map image, we propose to use the eight-neighborhood averaging method to determine the grayscale value of a pixel that has multiple correspondences. For example, if pixel coordinates (u_n, v_n) occur in both of the first two columns of matrix $M_{l_{oe}}$ and $M_{l_{ue}}$, we check the eight neighbourhood pixels $(u_n + i, v_n + j)$, where $i, j=1,0,-1$, find the average of the grayscale values that these pixel coordinates refer to, and assign the average value to pixel (u_n, v_n) given by $g_{u_nv_n}$. Then coordinates (u_n, v_n) are reallocated to the occupancy-coordinate matrix by comparing $g_{u_nv_n}$ with the threshold t_f and t_o , and are removed from previous matrices $M_{l_{oe}}$ and $M_{l_{ue}}$.

Moreover, if the scaling factor is larger than 1, there will exist unassigned pixels in the map image after scaling, which are regarded as holes. To maintain a consistent trend across the pixels, we propose to use a bilinear interpolation method to appropriately assign the grayscale values to these pixels by at least four well-assigned pixels. For example, the grayscale value of pixel (u_k, v_k) is unassigned, but the grayscale values at the pixels (u_1, v_1) , (u_1, v_2) , (u_2, v_1) and (u_2, v_2) are known. We first perform the linear interpolation in the u-coordinates as

$$\begin{aligned} g_{u_kv_1} &= \frac{u_2 - u_k}{u_2 - u_1} g_{u_1v_1} + \frac{u_k - u_1}{u_2 - u_1} g_{u_2v_1}, \\ g_{u_kv_2} &= \frac{u_2 - u_k}{u_2 - u_1} g_{u_1v_2} + \frac{u_k - u_1}{u_2 - u_1} g_{u_2v_2}, \end{aligned} \quad (5-18)$$

where $g_{u_1v_1}$, $g_{u_1v_2}$, $g_{u_2v_1}$ and $g_{u_2v_2}$ are the grayscale values of pixel (u_1, v_1) , (u_1, v_2) , (u_2, v_1) and (u_2, v_2) , respectively. Then we proceed by interpolating in the v-coordinates and substituting the results of $g_{u_kv_1}$ and $g_{u_kv_2}$ from (5-18) as

$$\begin{aligned} g_{u_kv_k} &= \frac{v_2 - v_k}{v_2 - v_1} g_{u_kv_1} + \frac{v_k - v_1}{v_2 - v_1} g_{u_kv_2} \\ &= \frac{1}{(u_2 - u_1)(v_2 - v_1)} [u_2 - u_k \quad u_k - u_1] \begin{bmatrix} g_{u_1v_1} & g_{u_1v_2} \\ g_{u_2v_1} & g_{u_2v_2} \end{bmatrix} \begin{bmatrix} u_2 - u_k \\ u_k - u_1 \end{bmatrix}. \end{aligned} \quad (5-19)$$

Using (5-19), each unassigned pixel can be determined by at least four pixels allocated with definite grayscale values. Thereby, after multiplying a scaling matrix, a complete map image can be obtained by a rounding operation, eliminating duplication, eight-neighborhood averaging and bilinear interpolation.

Furthermore, similarly to the calibration method for the map orientation described in Section 5.4.2, the calibration for scale is also scalable to the a perceived environment that contains multiple VLP landmarks, as long as we find the average of the scaling factor of every two key points on the map and substitute the average into (5-15).

5.5 Experimental Results

In this section, experiments are conducted to evaluate the performance of the proposed map calibration method. We will describe the experiment setup, evaluate map alignment performances and analyze the navigation results on the maps that are calibrated by the proposed method.

5.5.1 Experiment Setup

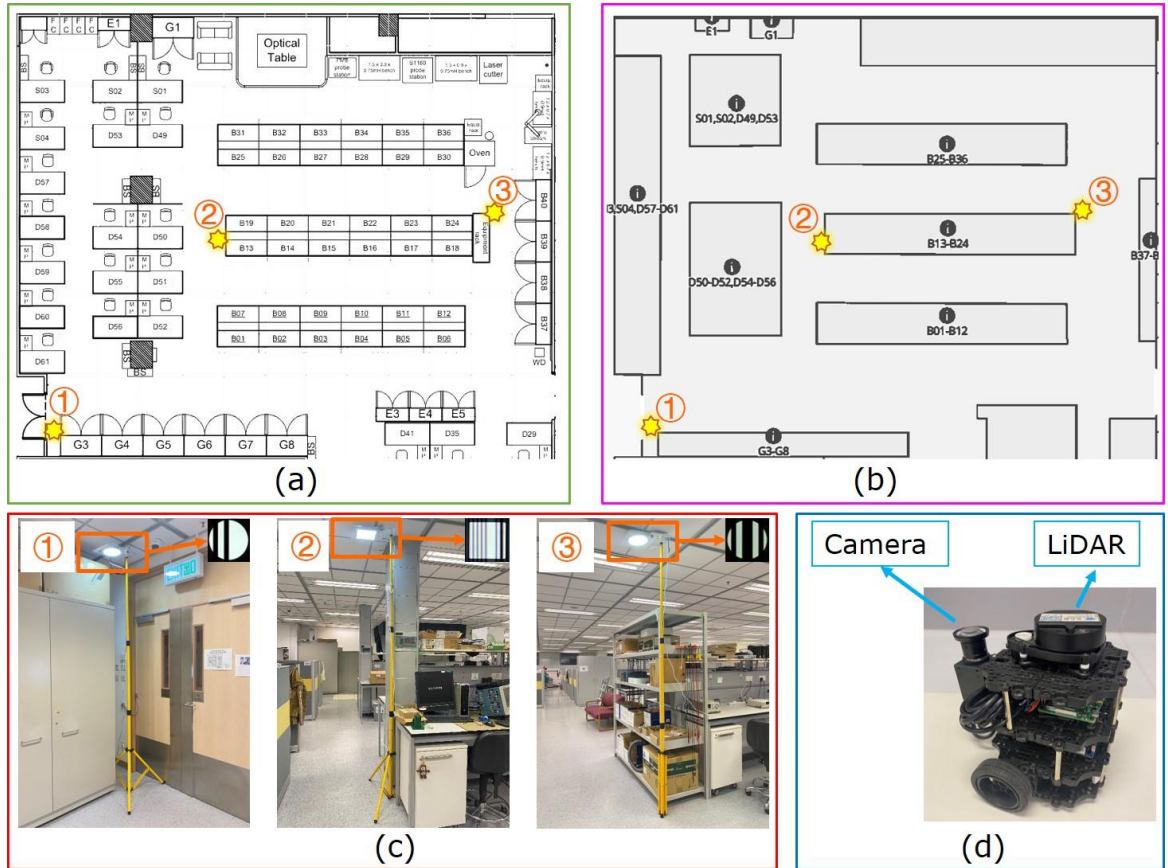


Figure 5.7: Experimental setup: (a) a building blueprint, (b) a floor map, (c) three VLP landmarks, (d) a TurtleBot3 Burger robot.

Table 5.1: Experimental Parameters

LED Height	2.7 m
LED Diameter	0.175 m
LED Power	18 w
Camera Resolution	2048×1536

The experiment is performed in our lab (Integrated Circuit Design Center, 3/F, CYT Building, HKUST). Two maps with different accuracies are prepared: one is a building blueprint with a high accuracy and the other is a floor map with a rough accuracy, as shown in Figure 5.7(a) and (b). We set three VLP lights as landmarks at different locations in the perceived environment, as illustrated in Figure 5.7(c). They are modulated with different IDs, which are encoded with different positioning information stored in the database. Figure 5.7(d) illustrates the TurtleBot3 Burger robot, the ROS standard platform robot we use for the VLP receiver and SLAM process. The robot is equipped with a Raspberry Pi 3 Model B, running Ubuntu 16.04 with ROS. Sensors are mounted on the robot for the SLAM process and VLP decoding. These include an IMU, odometer, 360° LiDAR and an industrial camera facing toward the ceiling. We use a laptop running Ubuntu 18.04 with ROS to remote control the robot and record the data from the robot in the mapping process. The experimental parameters and the camera options are summarized in Table 5.1.

5.5.2 Mapping Process and Alignment Results

In the mapping experiment, we set the robot under VLP light No. 1 as the starting point and control it to move to VLP light No.2 and then No.3. At the same time, the robot performs SLAM using the Gmapping [19] package in ROS. Thus, the robot’s location on the SLAM map and position as estimated by the VLP system are recorded synchronously, and the SLAM map is visualized in RViz software on the laptop, as shown in Figure 5.8. After the robot has tracked all three VLP lights, the occupancy grid map is saved. Then we perform the proposed map calibration method to calibrate the orientation and scale of the saved sensor map and the building blueprint. To intuitively evaluate the performance of the map calibration results, we align the two maps, specifically, translating the key points to the same pixels on the map. Then

we overlap the pixels which are occupied to compare the structures shown on different maps of the same experimental area.

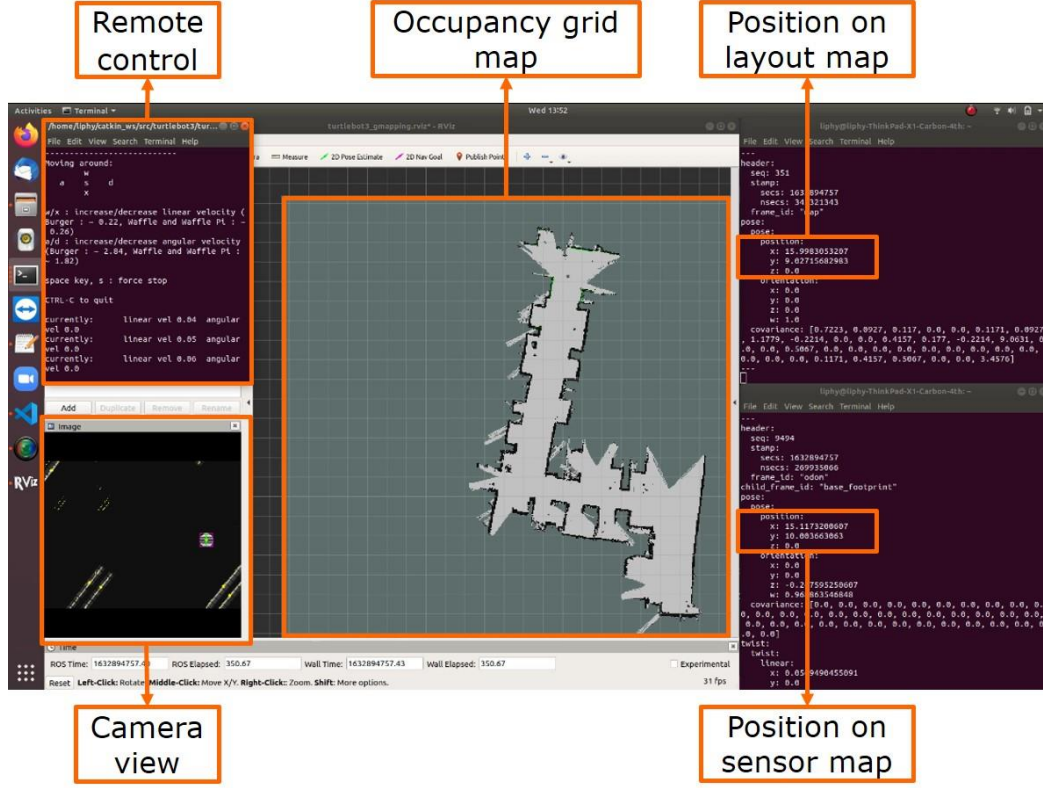


Figure 5.8: Mapping process visualized in RViz.

By multiplying the scaling matrix as (5-15) and rounding the values, the pixel coordinates of the key points on the layout map after scaling can be computed as $(\lfloor \frac{d_{u_l}}{d_{u_s}} u_{l1} + \frac{1}{2} \rfloor, \lfloor \frac{d_{v_l}}{d_{v_s}} v_{l1} + \frac{1}{2} \rfloor)$ and $(\lfloor \frac{d_{u_l}}{d_{u_s}} u_{l2} + \frac{1}{2} \rfloor, \lfloor \frac{d_{v_l}}{d_{v_s}} v_{l2} + \frac{1}{2} \rfloor)$. Then we translate the pixels on the sensor map so that the key points on the two maps are located at the same pixel coordinates. The translation matrix is given by

$$T_{s3} = \begin{bmatrix} 1 & 0 & \frac{\lfloor \frac{d_{u_l}}{d_{u_s}} u_{l1} + \frac{1}{2} \rfloor - u_{sm1} + \lfloor \frac{d_{u_l}}{d_{u_s}} u_{l2} + \frac{1}{2} \rfloor - u_{sm2}}{2} \\ 0 & 1 & \frac{\lfloor \frac{d_{v_l}}{d_{v_s}} v_{l1} + \frac{1}{2} \rfloor - v_{sm1} + \lfloor \frac{d_{v_l}}{d_{v_s}} v_{l2} + \frac{1}{2} \rfloor - v_{sm2}}{2} \\ 0 & 0 & 1 \end{bmatrix}. \quad (5-20)$$

Multiplying the translation matrix in (5-20), the translated matrices are given by

$$M_{s_{ot}} = T_{s3} M_{s_{om}}. \quad (5-21)$$

Then we find all the pixels on the layout map, whose coordinates are given in $M_{s_{ot}}$, and allocate these pixels with the grayscale values of the same pixels on the sensor map.

5.5.3 Navigation on Calibrated Map

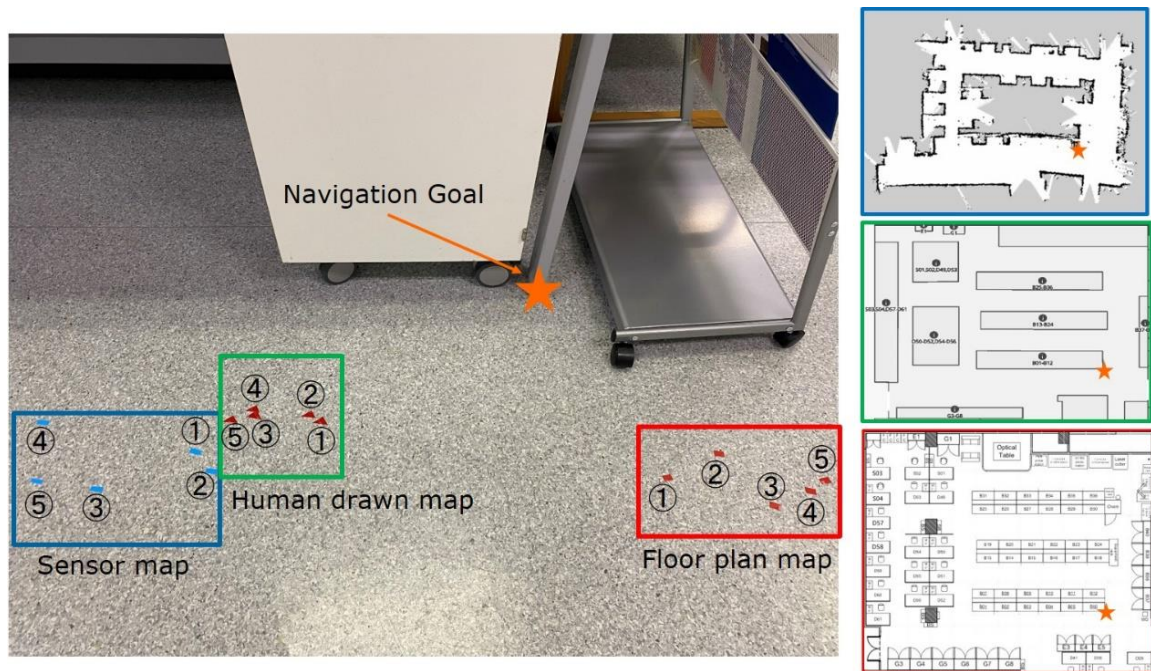


Figure 5.10: Navigation goal and navigation results on the sensor map, floor map and building blueprint.

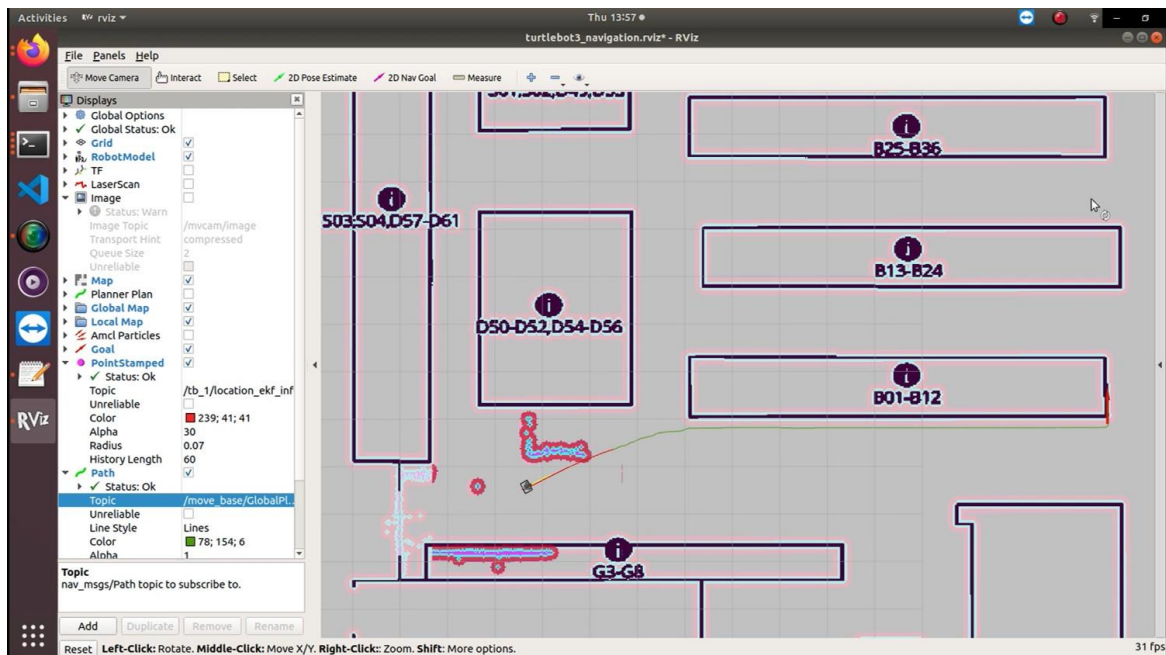


Figure 5.11: Navigation process with LiDAR detection results, DWA local plan and Dijkstra's global plan.

Table 5.2: Navigation Results

Points	Sensor Map (cm)	Building Blueprint (cm)	Floor Map (cm)
①	65	45	49
②	67	46	50
③	83	57	59
④	88	60	59
⑤	90	62	63
Average	78.6	54	56

To further evaluate the performance of the proposed map calibration method, we use a navigation system based on an adaptive Monte Carlo localization (AMCL) [20] package, Dijkstra’s algorithm [21] package and dynamic window approach (DWA) [22] package in ROS to achieve autonomous navigation of the robot on the calibrated maps. We set the navigation goal of the robot to be next to cubicle B06, as shown in Figure 5.10. The distance between the starting point and the target point is 14.14 m. During the navigation, the position of the robot is determined by the AMCL method. The global plan is achieved by Dijkstra’s algorithm, and the local plan is designed by the DWA planner. As shown in Figure 5.11, the blue dots encapsulated in the outlines in pink are the obstacles detected by the LiDAR on the robot. The red line is the DWA local planner, and the green line, which connects to the navigation goal, represented by a red arrow, is the global plan based on Dijkstra’s algorithm. On each map, we repeat the navigation five times and the navigation results on the different maps are illustrated in Figure 5.10 and summarized in Table 5.2. The table lists the distance between the actual point reached in the real world and the target destination sent to the robot. Compared with the navigation results on the sensor map, those on the calibrated building blueprint and the calibrated floor map are much closer to the set destination. The average navigation accuracy is improved by 24.6 cm on the building blueprint and 22.6 cm on the floor map, respectively. Furthermore, with the proposed calibration method, the robot on the floor map, which has lower accuracy in scale and structure location than the building blueprint, can achieve a navigation performance nearly as good as that on the building blueprint, which verifies the effectiveness

of the proposed map calibration method. As mentioned in Section 5.1, the sensor map generated using SLAM is with noise and distorted as shown in Figure 5.10. However, the noisy sensor map will not degrade the calibration performance as listed in Table 5.2 as the calibration process is based on the positions of the key points on the two different maps. The positions of the key points on the layout map are estimated by the proposed high-accuracy VLP system.

5.5.4 Navigation with Semantic Information

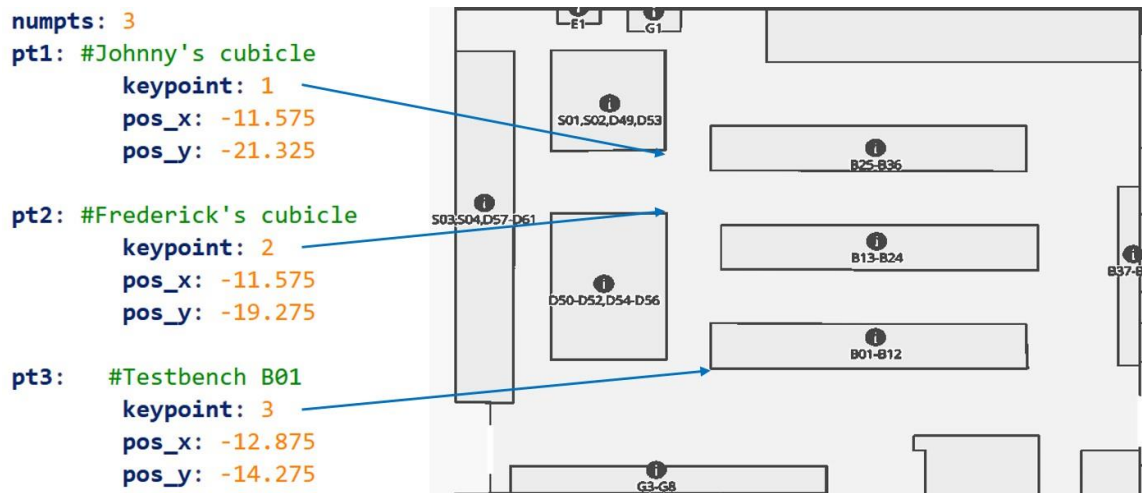
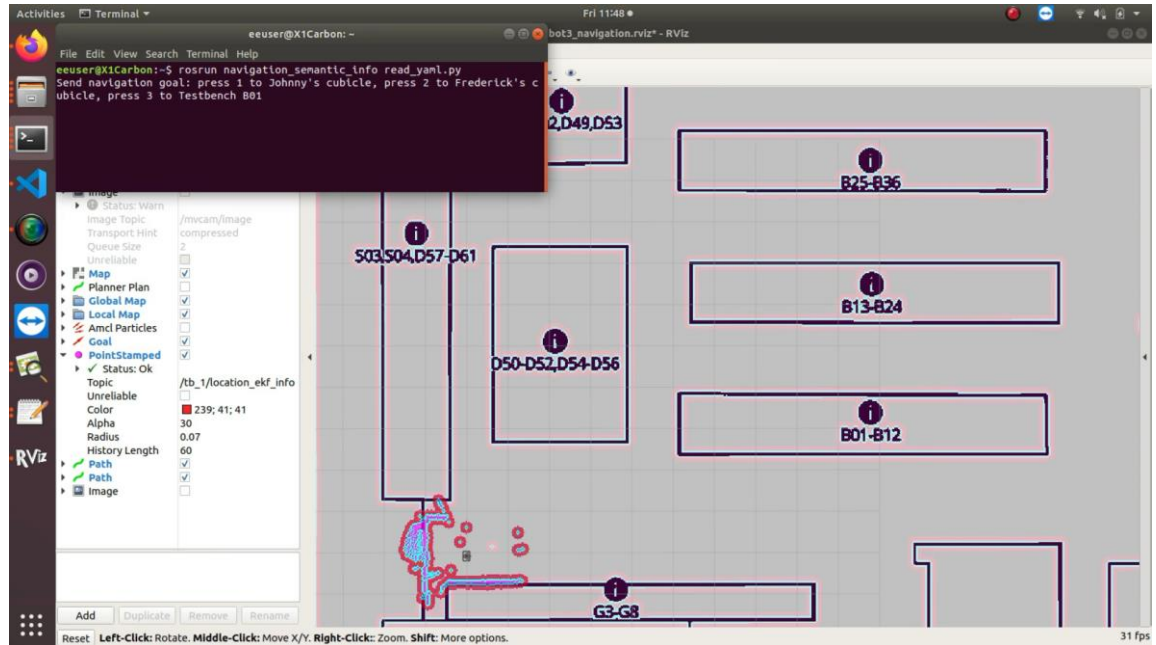


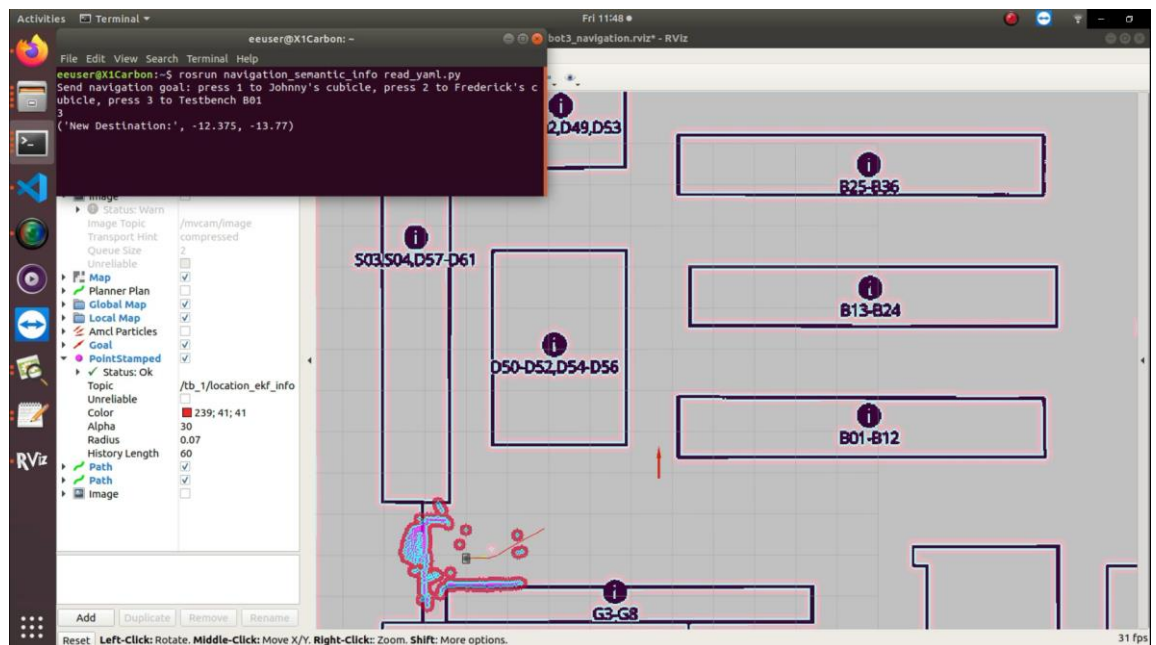
Figure 5.12: Compiled semantic information on the floor map.

As we mentioned in Section 5.1, a layout map contains semantic information which is accessible to humans and allows them to send instructions to robots. After calibrating the layout map, we cannot only send the navigation target to the robot by selecting one point on the map, but also navigate the robot with pre-compiled semantic information on the layout map. For the floor map illustrated in Figure 5.7(b), we compile three positions with semantic information, as shown in Figure 5.12, where D49 refers to Johnny's cubicle, D50 refers to Frederick's cubicle and B01 refers to the last line of the test bench. Figure 5.13 illustrates the navigation experiment with semantic information. When the program starts, the semantic information of the map is shown in the terminal, as shown in Figure 5.13(a). Then we send a navigation goal by tapping the target identifier number of the semantic information, and a red arrow indicating the aimed point is marked on the floor map, as shown in Figure 5.13(b). Figure 5.13(c) illustrates the condition that the robot has arrived at the target point and the semantic information is illustrated again in the terminal. We can repeat the navigation process by tapping

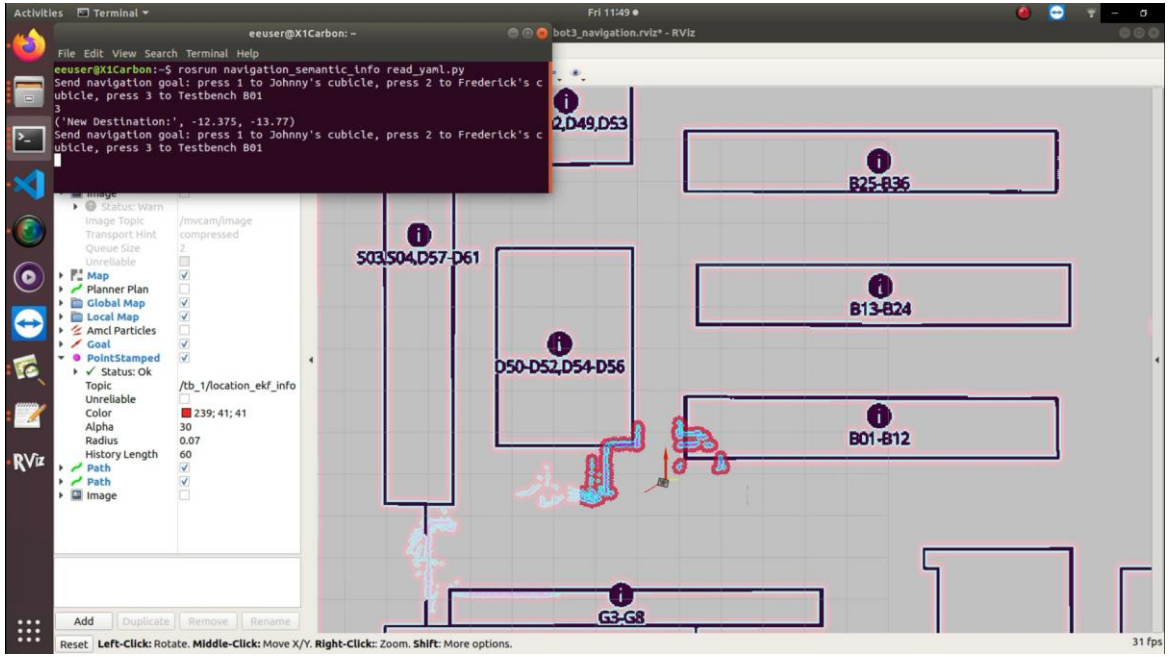
another identifier number of the navigation goal, as shown in Figure 5.13(d). Using this process, the semantic information on the floor map helps humans to set tasks for the robots in a more straightforward and user-friendly way compared with a sensor map, which has no semantic information.



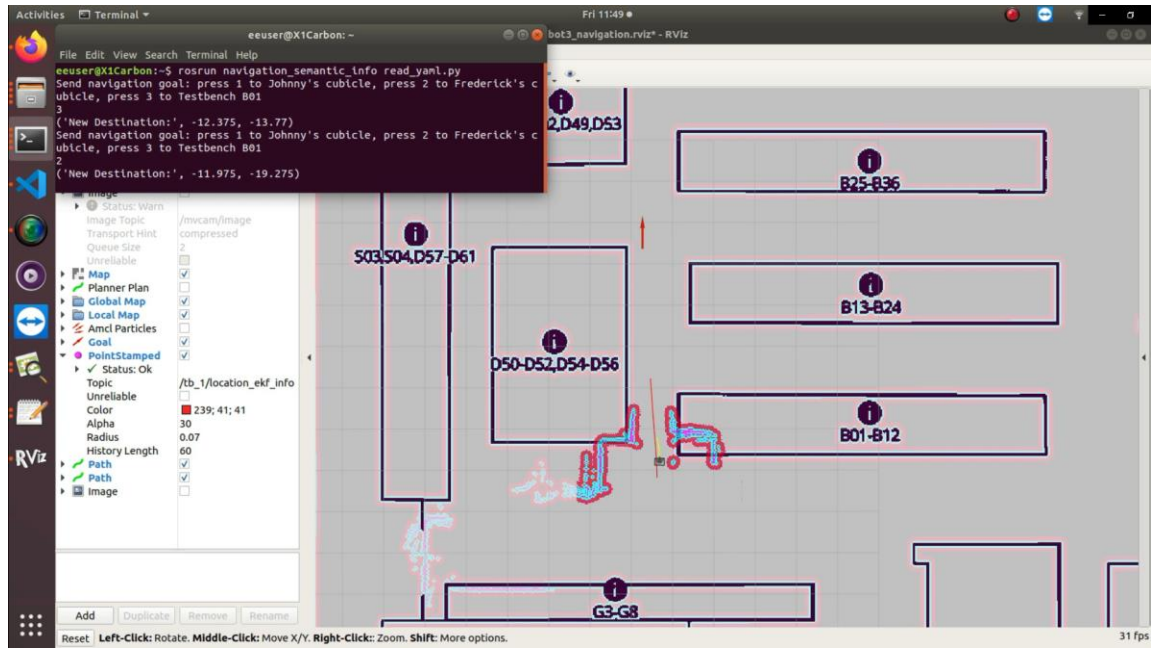
(a)



(b)



(c)



(d)

Figure 5.13: Navigation process with semantic information: (a) sending the command to the robot, (b) robot navigating to the target point, (c) after reaching the target point, robot waiting for the next navigation goal, (d) robot navigating to the second target point.

5.6 Summary

In this chapter, we propose a VLP landmark and SLAM-assisted automatic map calibration method for robot navigation. VLP landmarks with different IDs are mounted in the

environment to be perceived, and a layout map of the environment is prepared to be calibrated. By tracking the landmarks and conducting SLAM at the same time, the robot's position as obtained from the VLP method and the position as obtained from SLAM are recorded synchronously with the time stamp. By aligning the recorded coordinates on the layout map and the sensor map saved from SLAM, the orientation and the scale of the layout map is calibrated. Experiments are performed to evaluate the proposed map calibration system in terms of the map alignment performance and navigation performance. We calibrate two layout maps: a building blueprint of high accuracy and a floor map of rough accuracy. The experiment results show that the robot can achieve a better navigation performance on the calibrated layout maps compared with that on the sensor map, and can achieve the navigation performance on the calibrated floor map almost the same as that on the calibrated building blueprint.

5.7 References

- [1] 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. Tutor.*, vol. 20, no. 3, pp. 1963–1988, 2018.
- [2] W. Guan, L. Huang, B. Hussain, and C. P. Yue, "Robust robotic localization using visible light positioning and inertial fusion," *IEEE Sens. J.*, 2021.
- [3] Y. Wang, B. Hussain, and C. P. Yue, "Arbitrarily tilted receiver camera correction and partially blocked LED image compensation for indoor visible light positioning," *IEEE Sens. J.*, 2021.
- [4] H. Gao, X. Zhang, J. Yuan, J. Song, and Y. Fang, "A novel global localization approach based on structural unit encoding and multiple hypothesis tracking," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 11, pp. 4427–4442, 2019.
- [5] A. Huletski, D. Kartashov, and K. Krinkin, "VinySLAM: an indoor SLAM method for low-cost platforms based on the transferable belief model," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017, pp. 6770–6776.

- [6] C. Cadena, L. Carlone, H. Carrillo, Y. Latif, D. Scaramuzza, J. Neira, I. Reid, and J. J. Leonard, “Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age,” *IEEE Trans. Robot.*, vol. 32, no. 6, pp. 1309–1332, 2016.
- [7] T. M. Bonanni, B. Della Corte, and G. Grisetti, “3-D map merging on pose graphs,” *IEEE Robot. Autom. Lett.*, vol. 2, no. 2, pp. 1031–1038, 2017.
- [8] Z. Jiang, J. Zhu, Y. Li, J. Wang, Z. Li, and H. Lu, “Simultaneous merging multiple grid maps using the robust motion averaging,” *J. Intell. Robot. Syst.*, vol. 94, no. 3, pp. 655–668, 2019.
- [9] C. X. Guo, K. Sartipi, R. C. DuToit, G. A. Georgiou, R. Li, J. OLeary, E. D. Nerurkar, J. A. Hesch, and S. I. Roumeliotis, “Resource-aware large-scale cooperative three-dimensional mapping using multiple mobile devices,” *IEEE Trans. Robot.*, vol. 34, no. 5, pp. 1349–1369, 2018.
- [10] C. A. V. Hern´andez and F. A. P. Ortiz, “A real-time map merging strategy for robust collaborative reconstruction of unknown environments,” *Expert Syst. Appl.*, vol. 145, p. 113109, 2020.
- [11] F. Amigoni and A. Q. Li, “Comparing methods for merging redundant line segments in maps,” *Robot. Autom. Syst.*, vol. 99, pp. 135–147, 2018.
- [12] R. K. Ramachandran, Z. Kakish, and S. Berman, “Information correlated L´evy walk exploration and distributed mapping using a swarm of robots,” *IEEE Trans. Robot.*, vol. 36, no. 5, pp. 1422–1441, 2020.
- [13] D. Kakuma, S. Tsuchihiro, G. A. G. Ricardez, J. Takamatsu, and T. Ogasawara, “Alignment of occupancy grid and floor maps using graph matching,” in *2017 IEEE 11th international conference on semantic computing (ICSC)*. IEEE, 2017, pp. 57–60.
- [14] C. Georgiou, S. Anderson, and T. Dodd, “Constructing informative bayesian map priors: A multi-objective optimisation approach applied to indoor occupancy grid mapping,” *Int. J. Rob. Res.*, vol. 36, no. 3, pp. 274–291, 2017.

- [15]S. G. Shahbandi, M. Magnusson, and K. Iagnemma, “Nonlinear optimization of multimodal two-dimensional map alignment with application to prior knowledge transfer,” *IEEE Robot. Autom. Lett.*, vol. 3, no. 3, pp. 2040–2047, 2018.
- [16]J. Hou, H. Kuang, and S. Schwertfeger, “Fast 2D map matching based on area graphs,” in *2019 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE, 2019, pp. 1723–1729.
- [17]A. Elfes, “Using occupancy grids for mobile robot perception and navigation,” *Computer*, vol. 22, no. 6, pp. 46–57, 1989.
- [18]B. Hussain, C. Lau, and C. P. Yue, “Li-Fi based secure programmable QR code (LiQR),” in *JSAP-OSA Joint Symposia*. Fukuoka, Japan: Optical Society of America, 2017, p. 6p A409 6.
- [19]G. Grisetti, C. Stachniss, and W. Burgard, “Improved techniques for grid mapping with Rao-Blackwellized particle filters,” *IEEE Trans. Robot.*, vol. 23, no. 1, pp. 34–46, 2007.
- [20]S. Thrun, D. Fox, W. Burgard, and F. Dellaert, “Robust Monte Carlo localization for mobile robots,” *Artif. Intell.*, vol. 128, no. 1-2, pp. 99–141, 2001.
- [21]E.W. Dijkstra et al., “A note on two problems in connexion with graphs,” *Numer. Math.*, vol. 1, no. 1, pp. 269–271, 1959.
- [22]D. Fox, W. Burgard, and S. Thrun, “The dynamic window approach to collision avoidance,” *IEEE Robot. Autom. Mag.*, vol. 4, no. 1, pp. 23–33, 1997.

CHAPTER 6 Conclusion, Future Work and Publications

6.1 Summary of Contributions

In this thesis, a high-accuracy visible light positioning (VLP) system is proposed, based on which robotic navigation and map construction is also achieved. The key contributions can be divided into the following three parts:

(1) A tilted receiver camera correction and partially blocked LED image compensation method is proposed for indoor VLP systems. The proposed VLP system does not require the LEDs to be placed at high density and can eliminate the additional positioning errors caused by a tilted receiver camera in realistic scenarios. We propose two methods to detect the geometric features of the captured LED images, and the experimental results show that both methods perform better than the existing single LED-based VLP methods in ideal scenarios. Furthermore, the two proposed methods outperform the existing methods and can effectively suppress the performance degradation when an incomplete LED image is captured. Therefore, the proposed methods can provide stable positioning services and further improve the robustness of the VLP system.

(2) An indoor VLP platform for real-time robotic localization and navigation is developed. It implements a completed positioning system with LEDs as transmitter and a camera on the robot as receiver and the position information is shown on the remote terminate in real time. The proposed system design will stimulate a wide range of the innovative utilization of VLP technology. Based on the proposed robot positioning and navigation system, a panorama creation method is proposed which can generate a panorama at any target point using a robot mounted with a typical USB camera.

(3) A VLP landmark and SLAM-assisted automatic map construction method is proposed to improve robot navigation. VLP landmarks with different IDs are mounted in the environment to be perceived, and a layout map of the environment is prepared to be calibrated. By tracking the landmarks and conducting Simultaneous Localization and Mapping (SLAM) at the same time, the robot's position as obtained from the VLP method and the position as obtained from

SLAM are recorded synchronously with the time stamp. By aligning the recorded coordinates on the layout map and the sensor map saved from SLAM, the orientation and the scale of the layout map is calibrated. Experiments are performed to evaluate the proposed map calibration system in terms of the map alignment performance and navigation performance. We calibrate two layout maps: a building blueprint of high accuracy and a floor map of rough accuracy. The experiment results show that the robot can achieve a better navigation performance on the calibrated layout maps compared with that on the sensor map, and can achieve the navigation performance on the calibrated floor map almost the same as that on the calibrated building blueprint.

6.2 Future Work

The high-accuracy indoor positioning robotic system presented in the thesis will serve as a platform for further research in following key areas.

6.2.1 An Prior Information Assisted Distortion Elimination Method for Occupancy Grid Map Construction

An occupancy grid map contains three kinds of information representing by different values of gray scale given by occupied, free of scale and unknown. Each cell contains the probability that it is occupied. Occupancy grid maps are usually constructed by robots conducting SLAM. Mapping process requires the probability that cell is occupied. SLAM requires the probability that the cell is occupied and where the robot is. Conventional SLAM is based on light detection and ranging (LiDAR), odometer, inertial measurement unit (IMU) and other inertial sensors, and uses odometer for robotic positioning and LiDAR to detect obstacles and correct the accumulated error from odometer. However, one crucial drawback of SLAM is that the constructed map will get distorted with time due to the drift of the estimated robot position caused by the accumulated error from the sensors. By inducing VLP-based landmarks to a SLAM system to correct the distortion of the occupancy grid map caused by the accumulated error from the odometer, the absolute positioning results from VLP lights can help increase positioning accuracy. Furthermore, if we already have an accurate layout map, the gray scale

values on the layout map can provide the prior information for LiDAR detection and correct the distortion caused by the odometer as well. There are two main challenges in the proposed work. One is to remove the semantic information on a layout map as a layout map always contains semantic information for human readers. The other one is how to determine the weight of the probability provided by the layout map and the weight provided by the sensor observations and control input.

6.2.2 Automated 3D Reconstruction using Robot-mounted 360-Degree Camera with Visible Light Positioning Technology for Building Information Modelling Applications

Building Information Modelling (BIM) is a method using well-structured digital information for generating, delivering and handling data on a design project during its construction lifecycle. By creating a digital BIM, all necessary parties, including engineers and project coordinators, can communicate efficiently and effectively resulting in a higher overall value. BIM brings 3D construction information of every component in a building together, which helps to schedule, monitor and inspect the construction projects. Currently, BIM construction and project progress inspection is done by software engineers and staff holding 360-degree cameras to take photos at multiple locations. Although 360-degree photos captured at construction site are static and discontinuous with each other, they provide redundant visual information on the existing BIM model, which is the pain point of current solution. Therefore, facing the needs of a new generation of intelligent construction, how to construct a real-time dynamic self-reconfiguration high-accuracy positioning system based on the existing technology is of great significance. Through the combination of intelligent robots and omnidirectional vision, it is expected to provide BIM with full-cycle, multi-angle, high-precision environmental sampling data, so as to provide support for better optimization of the construction process.

We propose to use 360-degree photos to reconstruct 3D models for buildings and use indoor VLP technology for robots, such as automated guided vehicle (AGV), mounted with 360-degree camera to acquire and label the positioning data of the captured 360-degree photos. This helps to provide quick updates of the dynamic changes at the on-going construction site,

minimize the risk of mistakes, reduce discrepancies, and lower abortive costs. Furthermore, during the construction with BIM model, 360-degree camera-aided 3D reconstruction is an efficient way to survey the facilities and capture all of the information in 3D for lifecycle management.

The use of 360-degree cameras with indoor VLP technology to create 3D construction could become very useful to provide inspection of no matter how long and narrow spaces and add these spaces to BIM model. At present, the market has great demand for this technology, but there is still no suitable product available. On one hand, it is limited by the accuracy of indoor positioning, and on the other hand, it is limited by the accuracy of multi-view stereo. Therefore, the plan proposed in this project is intended to break through the bottleneck problem: 1) Utilize the combination of indoor mobile platform and VLP technology to achieve centimeter-level positioning capabilities in any indoor environment; 2) Use the 360-degree camera to dynamically collect environment images to realize the centimeter-level distributed environment reconstruction ability; So as to provide the first efficient solution for BIM industry applications.

Figure 6.1 describes the hardware and processing flow of the proposed project. VLC modulated LED lights are used as location beacons to provide location signals that are captured by the robots to calculate 3D position with cm-level accuracy that can be translated to determine the camera pose of the 360-degree camera mounted on top of the robots. Each 360-degree image is unfolded into four 2D images initially. As 3D reconstruction proceed, the number of dissected 2D images are optimized depending on the performance of our proposed visual-robotics-VLP-based 3D reconstruction scheme.

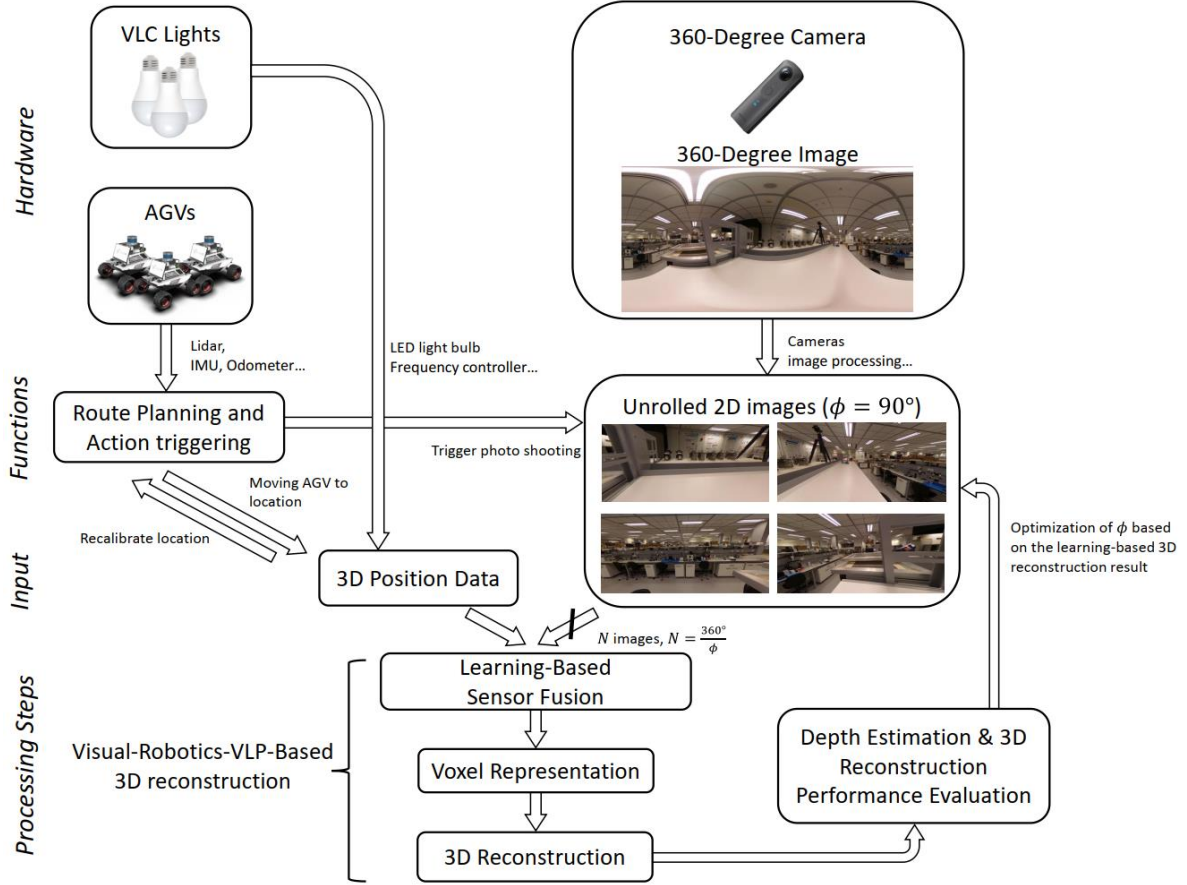


Figure 6.1: Conceptual diagram illustrating 3D reconstruction using 360-degree cameras mounted on multiple mobile robots, and visible light communication (VLC) sources used for 3D localization and camera pose estimation.

6.3 Publications

6.3.1 Works under review

- [1] **Y. Wang**, B. Hussain, and C. P. Yue, "VLP Landmark and SLAM assisted Automatic Map Calibration for Robotic Navigation," *ACM Transactions on Intelligent Systems and Technology*, 2022.
- [2] B. Hussain, **Y. Wang**, R. Chen, and C. Patrick Yue, "Camera Pose Estimation using a VLC-Modulated Single Rectangular LED for Indoor Positioning," *IEEE Transactions on Instrumentation & Measurement*, 2022.

- [3] B. Xu, B. Hussain, **Y. Wang**, H. Cheng, T. Min, and C. Patrick Yue, "Smart-Home Control System Using Visible Light Communication," *IEEE Transactions on Consumer Electronics*, 2022.

6.3.2 Published

- [1] **Y. Wang**, B. Hussain, and C. Patrick Yue, "Arbitrarily Tilted Receiver Camera Correction and Partially Blocked LED Image Compensation for Indoor Visible Light Positioning," *IEEE Sensors Journal*, vol. 22, no. 6, pp. 4800-4807, March 15, 2022.
- [2] **Y. Wang**, W. Guan, B. Hussain, and C. Patrick Yue, "High Precision Indoor Robot Localization Using VLC Enabled Smart Lighting," in *Proc. the 2021 Optical Fiber Communication Conference and Exhibition (OFC)*, June 06-11, 2021.
- [3] **Y. Wang** and C. Patrick Yue, "Orthogonally Interweaved Data Encryption Method for Screen to Camera Communication," in *Proc. the 35th Picture Coding Symposium (PCS2021)*, June 29-July 2, 2021.
- [4] **Y. Wang**, B. Xu, J. Kang, C. Qiu, and C. Patrick Yue, "Performance Analysis and Evaluation of Outdoor Visible Light Communication Reception," in *Proc. 5th EAI International Conference on Machine Learning and Intelligent Communications*, Shenzhen, China, September 26-27, 2020.
- [5] **Y. Wang**, L. Zhang, and Z. Wu, "Robust MUI Suppression for MIMO Visible Light Communication System with Location-Aided Chaotically Rotating Orthogonal Scheme," *IEEE Communications Letters*, vol. 23, no. 8, pp. 1361-1364, 2019.
- [6] **Y. Wang** and L. Zhang, "Uncoordinated Chaotic Channel Scrambling Scheme for Improving Security of MIMO-based VLC System," *IET Communications*, vol. 12, no.10, pp. 1245-1252, 2018.
- [7] **Y. Wang** and L. Zhang, "High Security Orthogonal Factorized Channel Scrambling Scheme with Location Information Embedded for MIMO-based VLC System," in *Proceedings of 2017 IEEE 85th Vehicular Technology Conference: VTC2017-Spring*, June 4-7, Sydney, Australia, 2017.

- [8] B. Hussain, **Y. Wang**, J. Cheng, R. Chen, and C. P. Yue, “LiDR: Visible Light Communication-Assisted Dead Reckoning for Accurate Indoor Localization,” *IEEE Internet of Things Journal*, 2022 (early access). DOI: 10.1109/IIOT.2022.3151664.