Stereo-Vision-Based Simultaneous Visible Light Communication and Range Estimation

Ruiyi Huang

Stereo-Vision-Based Simultaneous Visible Light Communication and Range Estimation

Ruiyi Huang

Graduate School of Engineering,

Nagoya University

2023

Contents

1 Introduction **1**

1.1 Background of Visible Light Communication 1

1.2 Research Trend of Visible Light Communication 4

1.3 Definitions of VLC receivers 6

1.3.1 Photo-Diode 7

1.3.2 Image Sensor 8

1.3.3 Camera 11

1.4 Image Sensor Communication (ISC) 12

1.4.1 Research Status of Image Sensor Communication 12

1.4.2 Image sensor communication and intelligent transportation system 17

1.4.3 Benefits and constraints of image sensor receivers 19

1.4.4 Simultaneous Image Sensor Communication and Ranging 22

1.5 Literature Review 23

1.5.1 Modulation Methods 23

1.5.2 Visible Light Ranging 25

1.5.3 Simultaneous Visible Light Communication and Ranging 26

1.6 Purpose of this study 30

1.7 Structure of this Dissertation 32

2 Vehicular Visible Light Communication and Ranging 35

2.1 Vehicle-to-everything communications using two high-speed cameras and LEDs 36

2.1.1 Exposure mechanism of high-speed cameras 38

2.1.2 Stereo Vision-Based Range Estimation 42

2.2 Simultaneous Ranging and Communication 43

2.3 Summary of Chapter 2 44

3 Simultaneous Visible Light Communication and Range Estimation Considering LED Saturation 45

3.1 System Model 46

3.1.1 Transmitter 46

3.1.2 Receiver 46

3.2 Bicubic Interpolation 49

3.3 Ranging 52

3.3.1 Phase-only Correlation 52

3.3.2 Sinc Function Matching 53

3.4 Communication 57

3.4.1 LED Luminance Extraction 57

3.4.2 Maximal ratio combining (MRC) 57

3.5 Experiment 58

3.6 Summary of Chapter 3 63

4 Visible Light Communication Using High Dynamic Range Images 65

4.1 Acquiring LEDs using High Dynamic Range (HDR) Settings 67

4.2 High Dynamic Range Combining 68

4.3 Experiment 70

4.3.1 Experiment setup 70

4.3.2 Luminance distribution 73

4.3.3 Performance Evaluation for PWM-8 78

4.3.4 Performance Evaluation for PWM-16 82

4.4 Summary of Chapter 4 84

5 Conclusion 87

5.1 Summary of this dissertation 87

5.2 Future Developments 88

Reference 91

List of Symbols 103

Research Publications 112

List of Figures

1.1 Number of research publications on visible light communication (VLC) shown by year [1,2], adopted from [3]. This result was obtained in February 2023 by searching Scopus for documents with publication year before 2023 and "visible light communication" within the title, keywords, and abstracts. 4

1.2 Categorizing the document type for VLC research publications. 5

1.3 The annual publication numbers of conference papers, article papers, and books were compared. Data was collected on July 2023, adopted from [3]. 6

1.4 Categorizing the receiver type for VLC research publications, adopted from [3]. 7

1.5 The cross section of a photo-diode 8

1.6 A 7×7 rolling shutter (RS) pixel array and its circuitry, adopted from [3]. The structure of one RS pixel is shown on the upper right, which has one floating diffusion (FD), one photodiode (PD), and 4 transistors, M_{RS}, M_{TG}, M_{SF}, and M_{SEL} [4]. 9

1.7 A 7×7 rolling shutter (RS) pixel array and its circuitry, adopted from [3]. The structure of one RS pixel is shown on the upper right, which has one floating diffusion (FD), one photodiode (PD), and 4 transistors, M_{RS}, M_{TG}, M_{SF}, and M_{SEL} [4]. 10

1.8 A camera structure 11

1.9 Number of research publications of image sensor communication (ISC) counted by year [5–7]. The result was obtained in February 2023 by searching the Scopus database for publications with publication year before 2023 and "image sensor communication" or "optical camera communication" or "screen camera communication" or "display camera communication" or {"visible light communication" and "camera"} or {"visible light communication" and "image sensor"} within the title, keywords, and abstracts. 13

1.10 Categorizing image sensor communication literature by applications. Others contains channel characterizing [8], modeling [9] and etc. 14

1.11 Categorizing literature related to image sensor communication by transmitter and receiver types. (a) Categorizing by transmitter type. Others include micro-LED [10], organic-LED [11], laser diodes, projector [12], etc. (b) Categorizing by receiver type. Others include charge-coupled device (CCD) camera, lensless-camera [13], time-of-flight sensor [14], etc. 16

1.12 Categories of various application scenarios and their proportions in the literature related to image sensor communication. Others include underground, etc. 18

1.13 This figure illustrates the spatial separation capabilities of an image sensor. When a car utilizes the camera to receive optical signals while driving, the lens will focus the light. Then the image sensor can effectively distinguish the light from the rear lights of the leading vehicle and the traffic signal based on their spatial locations. 20

2.1 The system transmits data signals from light sources such as traffic lights, vehicle headlights, and taillights. Two image sensors are installed on the vehicles to receive data and estimate the range simultaneously. The image sensor captures the visible light signals and converts them into electrical signals. These electrical signals are then decoded using image processing techniques. 37

2.2 The global-shutter exposure scheme is illustrated over time, wherein rows are simultaneously exposed and sequentially read out. 39

2.3 When a blinking LED array, as shown above, is captured by a global-shutter image sensor, the output image will exhibit the same as the first LED array pattern. It is assumed that the switching speed of the LED matches the readout rate of the sensor. 39

2.4 A diagram illustrating two calibrated cameras capturing a single LED is presented. The top view is displayed on the right-hand side. In this diagram, b denotes the distance between the left and right cameras, f signifies the camera's focal length, ρ represents the size of one pixel, P represents the LED, and Z indicates the distance that needs to be estimated. 41

3.1 Simultaneous visible light communication and range estimation system model. The upper part of the receiver is for the VLC data reception, and the lower part estimate the range. 47
3.2 The transmitter side and receiver side for PWM modulation with a modulation level of 4. 48
3.3 The process of receiving an LED and interpolate signal. $G(x, y)$ represents the received image and $G_{bic}(x_{bic}, y_{bic})$ represents the interpolation image. (x, y) and (x_{bic}, y_{bic}) are the corresponding points when $G(x, y)$ is interpolated into $G_{bic}(x_{bic}, y_{bic})$. 50
3.4 The POC function taken at a distance of 20 m, where the correlation peak is denoted by the color yellow. x and y are the horizontal and vertical axis of the image, respectively. z is the correlation value. 54
3.5 Phase correlation of different luminance level LEDs. Using sinc functions to match image pixels of low and high luminance LED has different characteristics. 56
3.6 Photograph showcasing the experimental setup. The LED transmitter is highlighted within a red rectangle. 59
3.7 Average Range Estimation Error 61
3.8 The symbol error rate is measured for OOK and PWM-8. 62

4.1 System Model for HDR combining. 66
4.2 A simplified channel model for HDR combining. 69
4.3 Experiment View. 71
4.4 Luminance distributions for PWM-8 at a distance of 20 m. The aperture values are (a) f/2. (b) f/4. (c) f/8. (d) f/16. 74

4.4 (*continued*) Luminance distributions for PWM-8 at a distance of 20 m. The aperture values are (a) f/2. (b) f/4. (c) f/8. (d) f/16. 75
4.5 Luminance distributions for PWM-16 at a distance of 20 m. The aperture values are (a) f/2. (b) f/4. (c) f/8. (d) f/16. 76
4.5 (*continued*) Luminance distributions for PWM-16 at a distance of 20 m. The aperture values are (a) f/2. (b) f/4. (c) f/8. (d) f/16. 77
4.6 Symbol error rate result for PWM-8. The left camera is f/4 for all the cases. The right camera is (a) f/8. (b)f/16. (c) f/4. 79
4.6 (*continued*)Symbol error rate for PWM-8. The left camera is f/4 for all the cases. The right camera is (a) f/8. (b)f/16. (c) f/4. 80
4.7 Throughput measurement result for PWM-8. 81
4.8 Symbol error rate for PWM-16. The left camera is f/4 for all the cases. The right camera is (a) f/8. (b)f/16. (c) f/4. 83
4.8 (*continued*)Symbol error rate for PWM-16. The left camera is f/4 for all the cases. The right camera is (a) f/8. (b)f/16. (c) f/4. 84
4.9 Throughput measurement result for PWM-16. 85

List of Tables

1.1 A summary of monocular ranging performance in the existing literature. 28
1.2 A summary of stereo ranging performance in the existing literature. . . . 29

3.1 Experiment parameters 60

4.1 Experiment parameters 72
4.2 Compare the minimal throughput of each reception method. 86

Chapter 1

Introduction

1.1 Background of Visible Light Communication

Visible light communication (VLC) is known as a wireless communication scheme that employs visible light to transmit data [2, 3, 15–17]. It utilizes light-emitting diodes (LEDs), laser diodes (LD), or liquid crystal displays (LCD) to emit visible light and transmit data. The VLC transmitters can be modulated extremely quickly to make it impossible for the human eye to detect the change in light [2].

VLC offers several advantages, such as security and compatibility with existing lighting infrastructure. It has applications in a variety of fields, including indoor positioning, smart homes, underwater communication, and intelligent transportation systems (ITSs). By harnessing the power of light, VLC enables high-speed data transmission, offering a promising alternative or complement to traditional wireless communication technologies. Moreover, the wavelength of visible light is significantly shorter than that of conventional radio waves. Therefore, VLC is perceived as a technology capable of solving the prob-

lems posed by the increasing demand for wireless communication spectrum [18].

The earliest recorded instances of VLC date back to the 7th century BC. In ancient China, elevated platforms were built at strategic locations or along key transportation routes during times of war. These platforms served as beacon stations that emitted smoke signals to notify people in distant places about the war. This was followed by a pioneering attempt by A. G. Bell to invent a "photophone" in the 1880s after entering the industrial age. Reference [19] describes this mechanical VLC system carried by electricity transmitting voice messages via sunlight. A. G. Bell considered this principle of optical communication using sunlight to be his greatest invention. However, early photophones did not gain widespread adoption in business owing to restrictions in communication distance and the unavoidable impact of barriers, such as adverse weather conditions. Nevertheless, photophones demonstrated the potential for VLC to be used for wireless communications.

The lighting industry underwent a revolutionary transformation with the emergence of high-brightness LEDs in the 1990s, as documented in the invention of high-brightness blue LEDs [20, 21]. This breakthrough not only opened up new avenues for VLC, but also introduced unprecedented possibilities. The remarkable responsiveness and rapid switching speed of LEDs enabled the transmission of high-speed data through visible light [2, 22]. Komine T. et al.'s (2004) study used white LED lights for both wireless communication and indoor illumination and demonstrated the potential of using LED lights to achieve high-speed communication at around 10 Gbps [2]. This paper [2] has greatly impacted the VLC field to be widely researched and explored.

VLC also finds applications outdoors, such as vehicle navigation and LED beacon communication. The first paper demonstrating the usage of an LED traffic light for communication can be found in [1] published in 1999, where a traffic light was encoded with

audio information. The receiver used a photo-diode (PD) to convert the light into an electrical signal through direct detection to demodulate the optically transmitted audio information. However, using a PD outdoors has several constraints owing to the presence of strong ambient noise. A novel type of receiver for VLC, image sensor (IS) receivers, was introduced in 2001. The IS receiver proved to be more robust to ambient noises and suitable for outdoor applications [5]. From 2015, the term "image sensor communication" (ISC) has been coined to refer to VLC systems that utilize ISs as receivers [23–25]. ISC is alternatively called "optical camera communication" (OCC) [26, 27].

Furthermore, some international standards of VLC have been established in the last decades. The IEEE 802.15.7-2011 standard was established in 2011 [28]. This standard provided a universal framework for short-range optical wireless communication utilizing visible light. It offers several key benefits, including access to a wide range of unlicensed spectrums spanning hundreds of terahertz. Additionally, it enhances and complements existing services provided by visible-light infrastructures, such as illumination, display, indication, decoration, etc. This standard [28] was subsequently modified to add ISC in 2018 as the IEEE 802.15.7-2018 standard [29]. Then, a new ISO standard was approved in 2020 for localized communication utilizing ISC in ITS stations [30]. This standard described a new communication port called "ITS-OCC", enabling the communication interface compatible with an ISC profile of IEEE 802.15.7-2018 for usage in ITS. With continuous enhancement of VLC standardization, VLC is expected to improve its performance and be applied to new cases, as well as further research and development.

1.2 Research Trend of Visible Light Communication

Figure 1.1 shows the annually published VLC literature before 2023 searched in the Scopus database. A total of 9,345 publications is found in the Scopus database for documents with "visible light communication" within the title, keywords, and abstracts, and publication year before 2023.

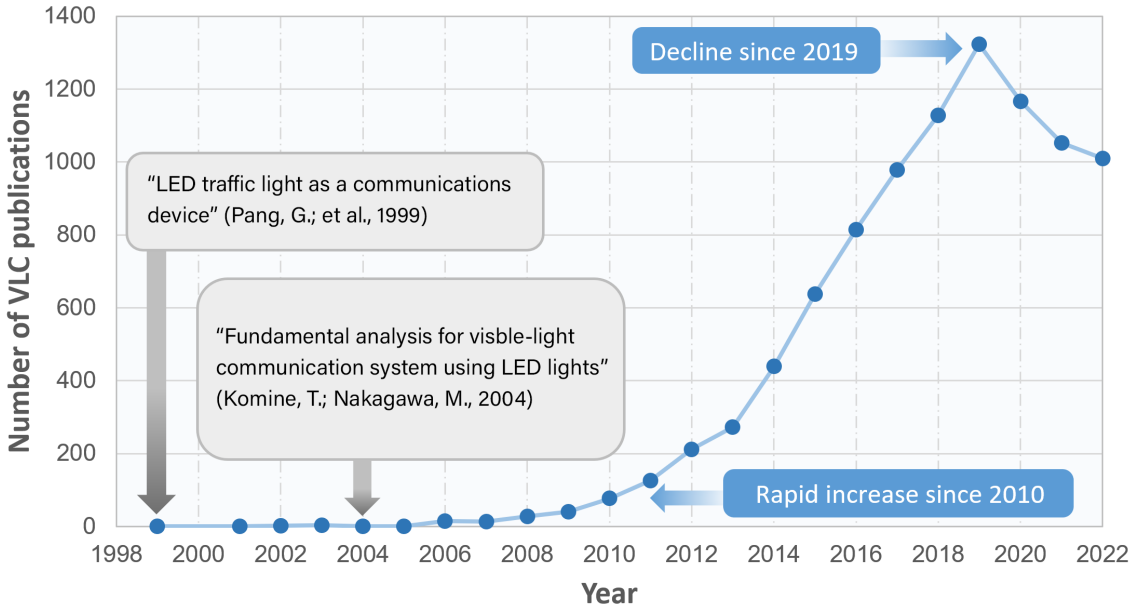


Figure 1.1: Number of research publications on visible light communication (VLC) shown by year [1, 2], adopted from [3]. This result was obtained in February 2023 by searching Scopus for documents with publication year before 2023 and "visible light communication" within the title, keywords, and abstracts.

Based on Figure 1.1, a notable increase in the number of VLC publications can be observed from approximately 2010, followed by a decline after 2019. The reason behind the decline can be inferred from Figure 1.2 and 1.3.

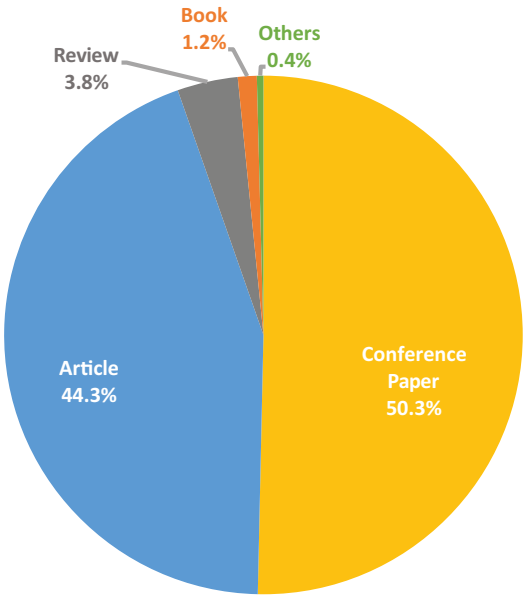


Figure 1.2: Categorizing the document type for VLC research publications.

Figure 1.2 presents the distribution of various document types in VLC literature. Notably, "conference papers" account for 50.3% of all publications. In Figure 1.3, the annual publication trends of "conference papers" and "articles and books" are investigated. The number of conference papers declined from 2019, while the articles and books category continue to increase. It can be inferred that the decline in VLC publications observed after 2019 is primarily due to the decrease in conference papers. The COVID-19 pandemic likely played a significant role in this decline, making it difficult for scholars to attend academic conferences starting in 2019. As a result, fewer conference papers were published during this period, leading to the overall decline in VLC publications.

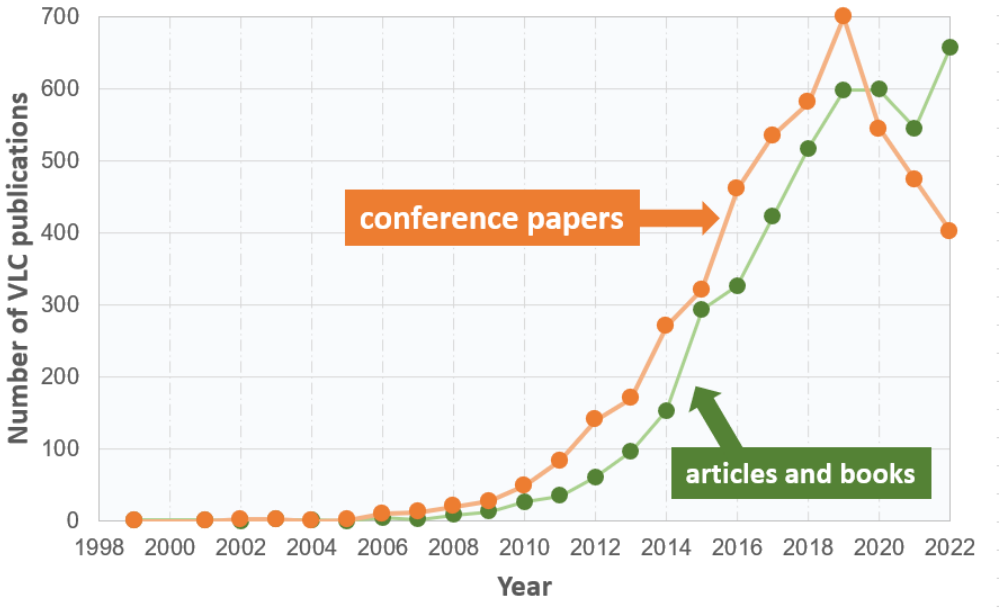


Figure 1.3: The annual publication numbers of conference papers, article papers, and books were compared. Data was collected on July 2023, adopted from [3].

1.3 Definitions of VLC receivers

Figure 1.4 shows the proportions of VLC research publications using PD or IS receivers. Among the publications, 88% employed PDs as the receiver, totaling 8,194 publications, while 12% utilized ISs, totaling 1,151 publications. Hence, it can be concluded that most existing works used PDs as the preferred receiver. Using PDs as the receiver offers several advantages, including rapid response and affordability. The applications of PD receivers include Light-fidelity (LiFi) [18] and indoor visible light positioning [31]. Additionally, the high-speed response of PDs resulting in communication at the Gbps level has been achieved [18,32,33]. In contrast, the communication speed of ISs is comparatively slower due to limitations in frame rates. This study focuses on VLC using IS receivers, aiming

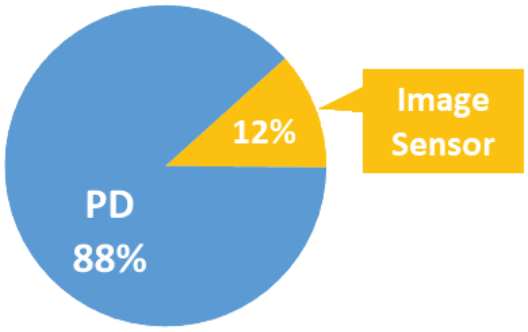


Figure 1.4: Categorizing the receiver type for VLC research publications, adopted from [3].

to solve the challenges of low frame rates.

1.3.1 Photo-Diode

A PD is a semiconductor device with a P-N junction that converts photons (or light) into electrical current, as shown in Figure 1.5. The P-layer has an abundance of holes (positive), and the N-layer has an abundance of electrons (negative). PDs can be fabricated using various materials, such as Silicon, Germanium, and Indium Gallium Arsenide. The selection of materials depends on specific requirements, including cost considerations, enhanced sensitivity, extended wavelength range, reduced noise levels, or faster response speeds. A single PD used in VLC usually involves directly converting the optical signals to electrical signals. It is an analog process.

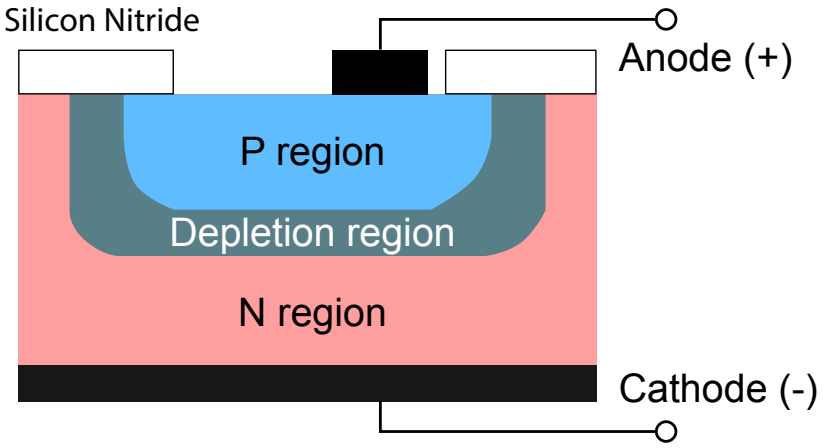


Figure 1.5: The cross section of a photo-diode

1.3.2 Image Sensor

Unlike a single PD receiver, IS is an electric device that detects and processes information to create an image. An IS comprises an extensive array of PD circuitry known as pixels. IS captures images by converting the varying attenuation of light waves, as they interact with objects, into electrical signals, represented as small bursts of current, to convey the visual data. In Figure 1.6, a typical four-transistor rolling shutter IS is illustrated. Row selectors scan the rows from top to bottom sequentially, followed by column selectors that read out the entire row. Due to the lack of a memory node, the pixels in the rolling-shutter sensor cannot store the accumulated charge. Therefore, readouts should be taken immediately after the exposure to ensure consistent exposure times for each row. The circuit diagram of a conventional five-transistor global shutter IS is illustrated in Figure 1.7. Similar to the rolling shutter, this sensor reads out data row by row. However, the global shutter

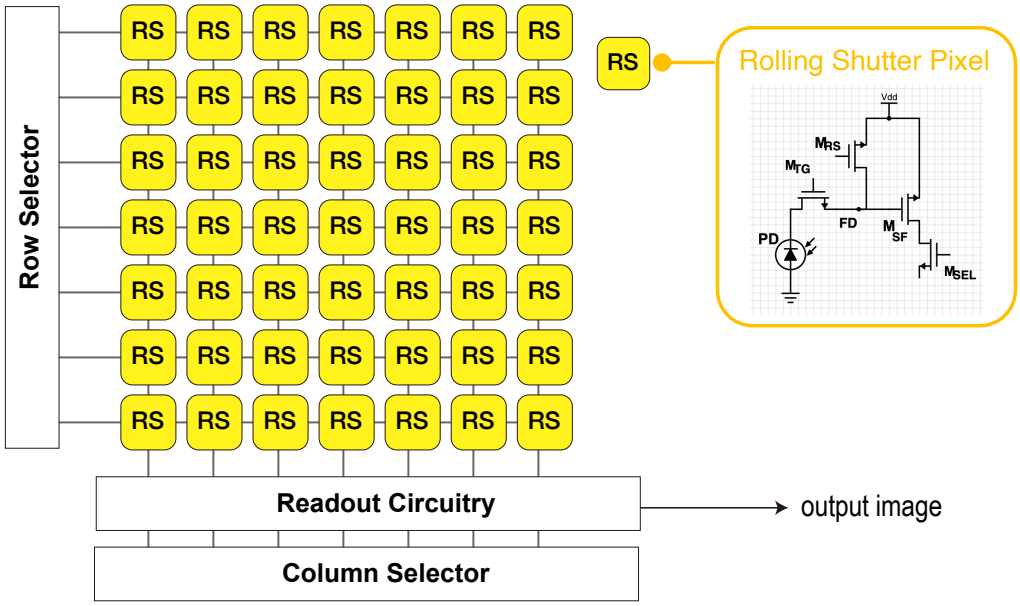


Figure 1.6: A 7×7 rolling shutter (RS) pixel array and its circuitry, adopted from [3]. The structure of one RS pixel is shown on the upper right, which has one floating diffusion (FD), one photodiode (PD), and 4 transistors, M_{RS}, M_{TG}, M_{SF}, and M_{SEL} [4].

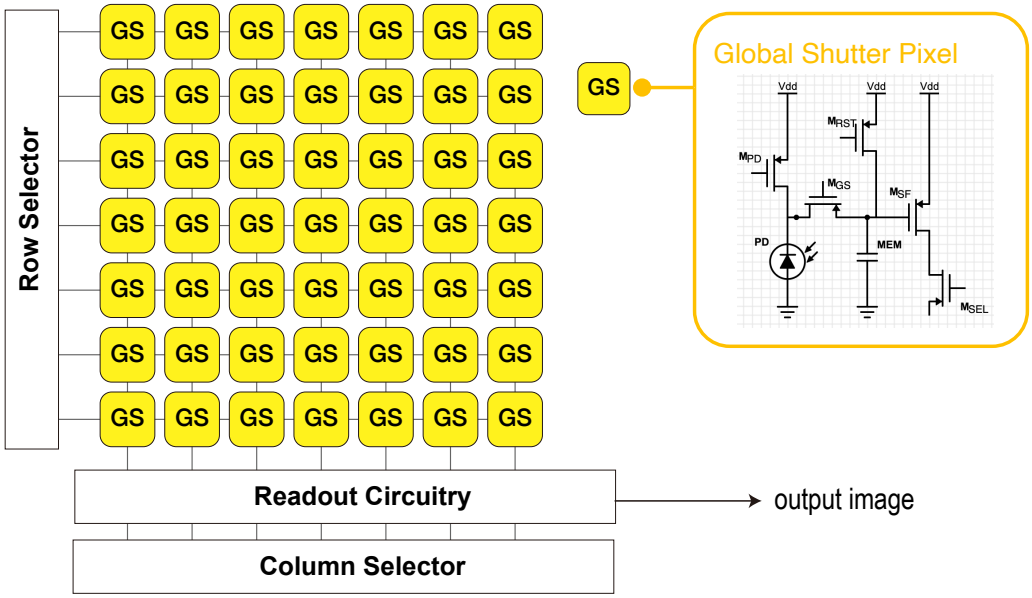


Figure 1.7: A 7×7 rolling shutter (RS) pixel array and its circuitry, adopted from [3]. The structure of one RS pixel is shown on the upper right, which has one floating diffusion (FD), one photodiode (PD), and 4 transistors, M_{RS}, M_{TG}, M_{SF}, and M_{SEL} [4].

pixel incorporates a memory node, allowing it to store accumulated charges. Therefore, a global-shutter IS can avoid image skewing and capture high-speed motions.

1.3.3 Camera

In this study, the term “camera” refers specifically to digital cameras instead of film cameras. A camera is defined as equipment consisting of an IS, a lens group, and a camera body, as illustrated in Figure 1.8. The IS is an electrical device component of the camera system.

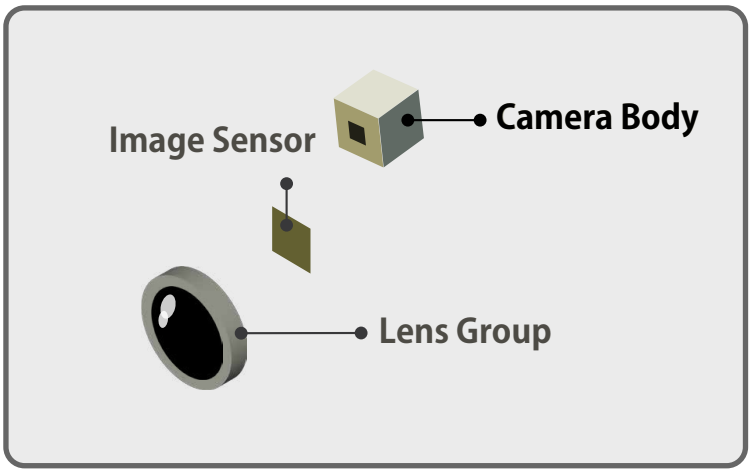


Figure 1.8: A camera structure

1.4 Image Sensor Communication (ISC)

This section introduces the research status and applications of ISC. In terms of the IS receiver, it is worth noting that an IS comprises a vast array of PDs. This unique structure enables the IS to spatially differentiate light sources, making it highly robust against outdoor ambient noise and facilitating accurate tracking of moving vehicles. ISs are particularly well-suited for outdoor applications, outperforming PDs in various aspects. One key advantage is their superior robustness to solar noise compared to PDs. Additionally, ISs offer better capabilities for tracking moving vehicles, which can pose challenges for PD-based systems. In outdoor environments, especially in Intelligent Transportation Systems (ITS), the impact of background light noise, such as sunlight and other light sources, cannot be overlooked. Consequently, the receiver's viewing angle must be narrowed to maintain an adequate signal-to-noise ratio. Moreover, mechanical adjustments are often necessary to align the receiver with the transmitter's direction. However, when a vehicle travels at high speeds, it becomes challenging for a PD with a restricted viewing angle to adapt its optical axis accordingly. In contrast, an IS can spatially distinguish between light sources, selectively capturing only the desired transmitted signal for communication while disregarding pixels that capture background light noise, such as sunlight.

1.4.1 Research Status of Image Sensor Communication

The chart presented in Figure 1.9 illustrates the quantity of ISC research literature published until 2023, as obtained from a search in Scopus. The total count amounts to 1,151 papers. Notably, Figure 1.9 demonstrates a substantial increase in publication numbers starting from 2013, followed by a plateau in growth after 2019.

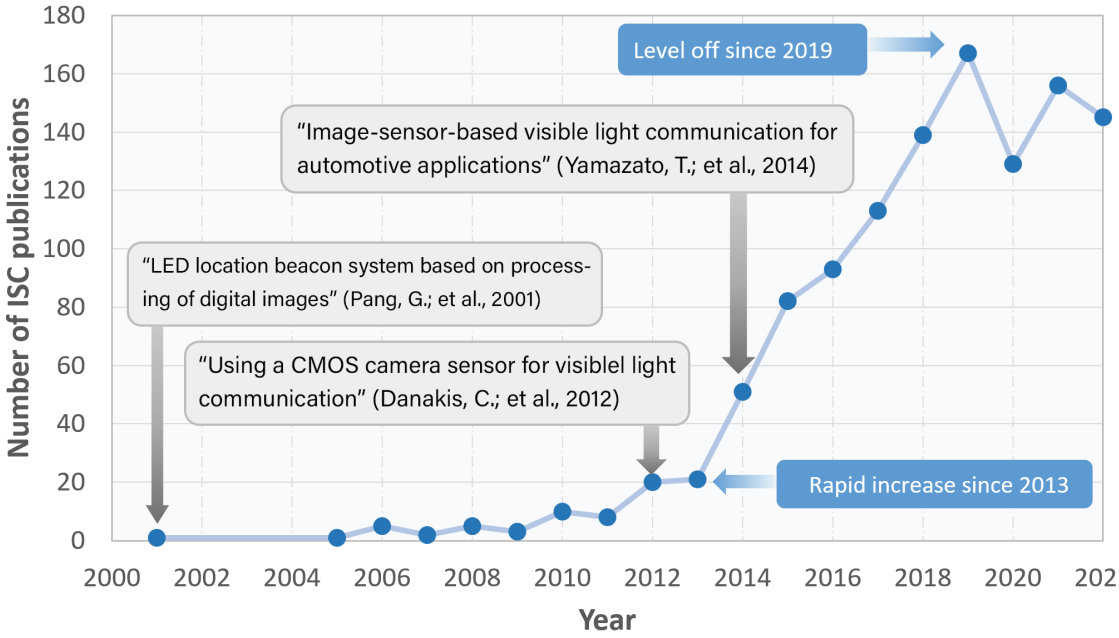


Figure 1.9: Number of research publications of image sensor communication (ISC) counted by year [5–7]. The result was obtained in February 2023 by searching the Scopus database for publications with publication year before 2023 and "image sensor communication" or "optical camera communication" or "screen camera communication" or "display camera communication" or {"visible light communication" and "camera"} or {"visible light communication" and "image sensor"} within the title, keywords, and abstracts.

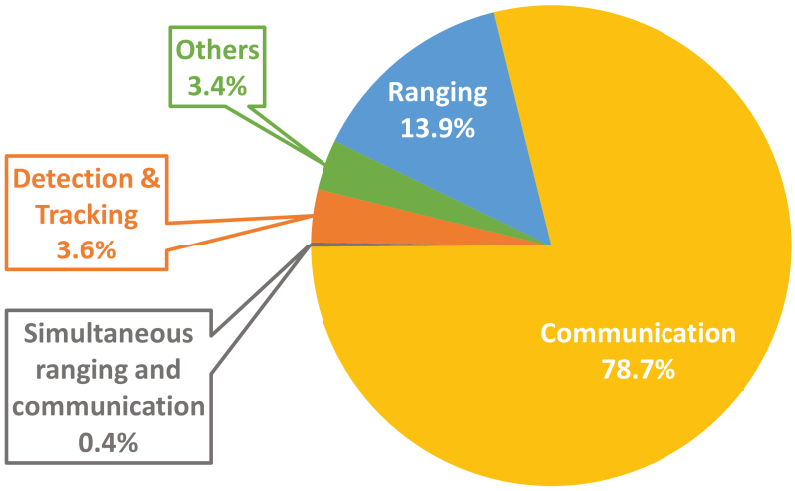


Figure 1.10: Categorizing image sensor communication literature by applications. Others contains channel characterizing [8], modeling [9] and etc.

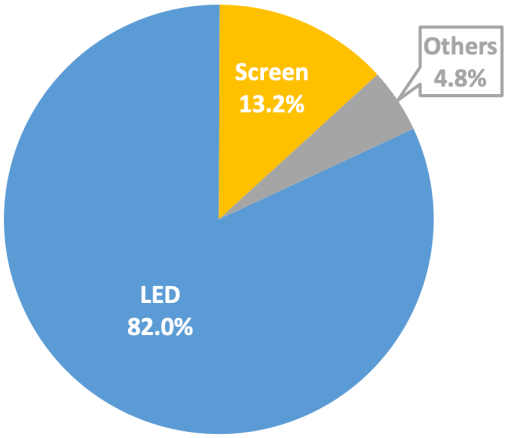
In order to analyze the research status of ISC further, the research publications with document types "article" and "conference paper" were selected and subsequently organized according to their applications (Figure 1.10), transmitter types (Figure 1.11(a)), and receiver types (Figure 1.11(b)). It is important to note that the categorization process was performed manually, which introduces the possibility of subjective interpretation and potential errors in the classification.

Figure 1.10 shows different applications of ISC. Communication occupies 78.7% of the overall publications. This category refers to transmitting information through light signals, such as data, voice, and video. This category encompasses various topics, including data rate measurement, communication model design, etc. The ranging application occupies 13.9% of the whole. Ranging, also known as localization, is an important aspect

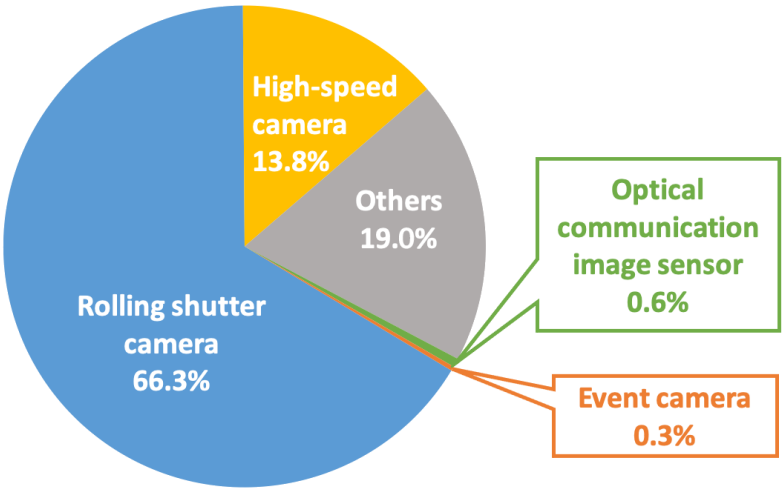
of ISC that finds applications in obstacle avoidance and navigation. Ranging in ISC relies on computer vision techniques. The positioning technology based on ISC utilizes LEDs for both positioning services and illumination, resulting in minimal additional power consumption and environmental friendliness. This approach reduces the upfront costs and eliminates the additional installation of VLC transmitters, thereby reducing the complexity and expenses associated with outdoor positioning systems. By employing ISC, vehicle communication, and high-accuracy localization can be achieved within a remarkable timeframe of 2 milliseconds using high-speed cameras, surpassing the capabilities of light detection and ranging or the global positioning system [34]. Moreover, tracking and detection comprise 3.6% of the entire ISC literature. This component plays a crucial role in both communication and ranging applications. It is imperative to achieve accurate detection, as any inaccuracies in tracking can result in erroneous data. Methods like machine learning can be employed for effective tracking purposes.

Figure 1.11 illustrates the distributions of transmitters and receivers in ISC. However, it should be noted that Figure 1.11(b) might exhibit a certain level of imprecision. This is primarily due to the absence of explicit information regarding the type of shutter used in some literature. Consequently, these articles are classified under the “others” category. Owing to the time limitations, the “others” category could not be comprehensively classified. It is possible that certain instances of rolling shutter cameras were probably erroneously classified as “others” in the analysis.

The category of transmitters in ISC encompasses two or more types. Figure 1.11 illustrates that LEDs are utilized significantly more than screens. LEDs prove to be highly suitable transmitters for illumination applications. In addition to being incorporated into existing lights, LEDs offer the advantage of rapid switching speeds. LED transmitters can



(a)



(b)

Figure 1.11: Categorizing literature related to image sensor communication by transmitter and receiver types. (a) Categorizing by transmitter type. Others include micro-LED [10], organic-LED [11], laser diodes, projector [12], etc. (b) Categorizing by receiver type. Others include charge-coupled device (CCD) camera, lensless-camera [13], time-of-flight sensor [14], etc.

further be classified into various subcategories, including single LEDs, fix LED arrays, and rotating LED arrays. On the other hand, the receiver category of ISC is divided into more than four types, as demonstrated in Figure 1.11(b). These include rolling shutter cameras, global shutter cameras, event cameras, and optical communication image sensors (OCIs).

1.4.2 Image sensor communication and intelligent transportation system

Figure 1.12 provides insights into the distribution of ISC across various application scenarios. These applications can be categorized into more than three types: outdoor, indoor, and underwater. Among these applications, the outdoor applications of ISC play a particularly crucial role in driving its widespread adoption. Outdoor applications of ISC hold immense potential for revolutionizing transportation systems and enabling advanced functionalities. Within the realm of ITS, ISC offers numerous benefits, such as improved safety measures, enhanced traffic flow optimization, and effective communication for emergency vehicle alerts. By harnessing the power of light signals, ISC can empower vehicles to achieve high-accuracy positioning and facilitate seamless communication within milliseconds. This study focuses on exploring outdoor applications, specifically focusing on ITS.

ITS encompass the integration of information and communication technologies to establish a connected framework between individuals, road infrastructure, and vehicles. The primary objective of ITS is to enhance road traffic safety, optimize traffic flow by mitigating congestion, and achieve overall efficiency and convenience in transportation. One notable technology utilized in ITS is dedicated short-range communications

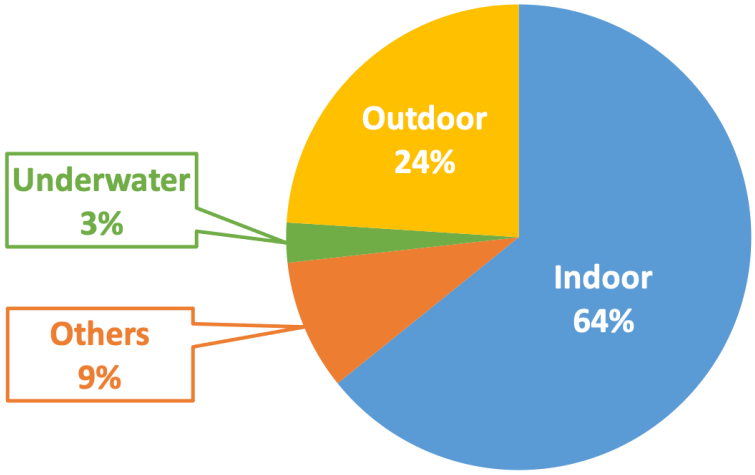


Figure 1.12: Categories of various application scenarios and their proportions in the literature related to image sensor communication. Others include underground, etc.

(DSRC) [35, 36], which enables secure and efficient wireless communication between vehicles and roadside infrastructure. This communication system plays a pivotal role in enabling various ITS applications. An example of an ITS application that many are familiar with is the electronic toll collection (ETC) system, which streamlines the toll payment process and reduces congestion at toll booths. Currently, the advancement of autonomous vehicles is garnering significant attention worldwide, particularly within the realm of ITS. The development of autonomous vehicles relies heavily on the ability to perceive and understand the surrounding road conditions using a variety of in-vehicle sensors. These sensors ensure driving safety and assist in autonomous driving scenarios. By leveraging advanced sensor technologies and communication systems, ITS plays a vital role in

shaping the future of transportation and revolutionizing the way we travel.

Moreover, the utilization of ISs within vehicles is on the rise, as indicated by the increasing number of ISs per vehicle [37]. These in-vehicle ISs serve various purposes, such as providing visual aids, facilitating object detection, and functioning as driving recorders. Furthermore, it is anticipated that the performance of these ISs, including image quality, viewing angle, and capture speed, will continue to improve over time. When combined with LED light sources, these in-vehicle ISs can also be employed as receivers for VLC, thereby augmenting the communication capabilities of vehicles. Numerous LEDs on the road, such as traffic signals, road signs, road studs, vehicle headlights, and taillights, present ample opportunities for leveraging VLC technology. In the context of autonomous vehicles, effective information exchange between roadside devices and vehicles is crucial for enabling safe driving and providing driving support. This encompasses various forms of communication, including roadside-to-vehicle and vehicle-to-vehicle interactions. By harnessing the capabilities of ISC, the exchange of information between vehicles and the surrounding infrastructure can be facilitated, leading to enhanced traffic control, improved road safety measures, and advancements in autonomous driving technologies.

1.4.3 Benefits and constraints of image sensor receivers

Spatial separation is a crucial characteristic of IS receivers in ISC. It involves the separation of signals originating from varied transmitters. Figure 1.13 illustrates the spatial separation of a traffic light and the leading vehicle, highlighting the process of separating and distinguishing these signals in the captured image.

ISC enables parallel transmission through spatial separation, allowing the utilization

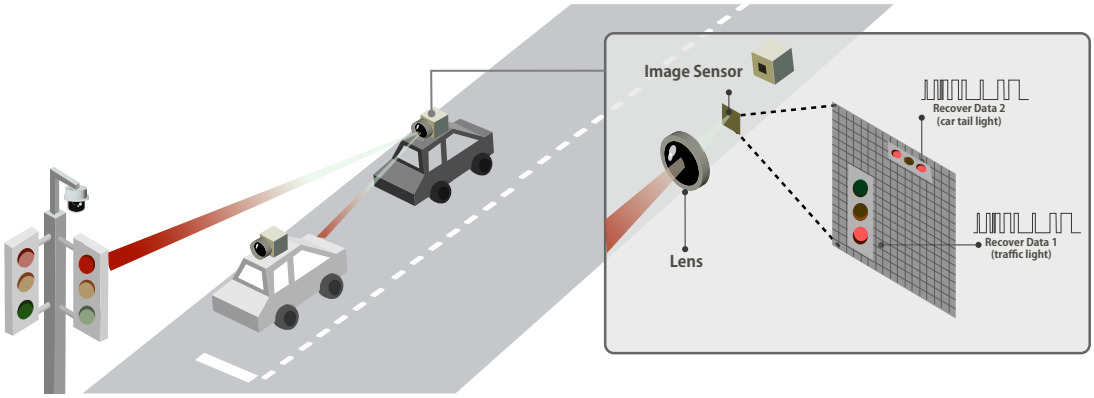


Figure 1.13: This figure illustrates the spatial separation capabilities of an image sensor. When a car utilizes the camera to receive optical signals while driving, the lens will focus the light. Then the image sensor can effectively distinguish the light from the rear lights of the leading vehicle and the traffic signal based on their spatial locations.

of multiple light sources that can be separated by image processing. Spatial separation is critical to minimize interference between these light sources, which would otherwise be challenging to separate using a single PD. To achieve spatial separation, LED arrays are commonly used in ISC systems [9, 38]. LED arrays have multiple LEDs arranged in a regular pattern, with each LED element individually controllable to illuminate at varying intensities. The data can be transmitted to the receiver by encoding it into an optical signal through the modulation of light generated by the LEDs in the array. The usage of LED arrays facilitates parallel transmission and greatly increases the data rate. Even minor interference can cause major errors in high-speed data transmission. Using IS receivers helps mitigate interference problems when transmitting multiple optical signals, enabling more efficient parallel transmission.

Moreover, the spatial separation capabilities of IS prove advantageous in distance

measurement applications. IS can achieve accurate object detection by providing depth information through a wide field of view using monocular and stereo vision. The depth data is invaluable to the vehicle as it aids in making safe and precise driving decisions.

Three key parameters of the IS influence the performance of ISC: resolution, dynamic range, and frame rate.

Firstly, the IS's resolution determines the level of detail captured by the transmitters. ISs with low resolutions capture images through fewer pixels, limiting the amount of detail that can be obtained compared to higher-resolution sensors. This can result in interference between adjacent transmitters, impacting the accuracy of image processing algorithms. Additionally, low-resolution images contain less data, leading to slower transmission rates, reduced reliability, and increased vulnerability to errors. Addressing the issue of low resolution can be achieved through the use of expensive high-resolution cameras or more affordable approaches using image processing techniques such as those proposed in [39] and [40].

Secondly, ISC faces limitations due to the low dynamic range of ISs. When photons strike the sensor plane, they are converted into electrons and accumulate in the pixel potential well. However, there is a limit to the number of electrons that can be accumulated, known as the saturation electron capacity. Once this limit is reached, no further photoelectric conversion is possible, resulting in a saturated output signal. The solution to low dynamic range is the main focus of this study.

Finally, the frame rate of the IS is an important factor to consider. Lower frame rates can lead to a decrease in the data rate. Conventional ISs typically operate at around 30 fps [4]. While there are commercial cameras available that can exceed this frame rate, such as the iPhone 14, which supports slo-mo videos at 120 fps or 240 fps [41], it is

important to note that achieving higher frame rates may come at the expense of video quality.

1.4.4 Simultaneous Image Sensor Communication and Ranging

In the context of ITS, the significance of low latency and accurate localization cannot be overstated [35, 42, 43]. Low latency enables reliable communication between vehicles and various transportation infrastructure components. Reliable coordination is crucial for optimizing traffic flow, preventing accidents, and ensuring timely responses. Precise localization is vital for vehicle navigation and vehicle collision avoidance.

ISC meets the requirements of low latency and accurate localization. On the one hand, it can achieve low latency by leveraging specific high-speed digital ISs. It is important to enable real-time data transfer between vehicles, road infrastructure, and other elements within the ITS ecosystem [7]. On the other hand, ISC systems harness the inherent properties of visible light, particularly its high brightness, to achieve high localization accuracy. ISC algorithms can effectively distinguish LED transmitters from dark or complex backgrounds, even in challenging traffic environments such as tunnels or intersections. In contrast, conventional wireless technologies often encounter difficulties related to signal reflection, multi-path propagation, and environmental interference, leading to inaccurate localization [44].

Furthermore, ISC and ranging can be made more efficient and simplified by performing communication and ranging functions simultaneously. There are several advantages to integrating communication and ranging into a single device. One significant benefit is the ability to use ranging information for data demodulation and also to use the demodulated data for ranging purpose. This interrelationship between communications and

ranging promotes mutual improvement and increases overall performance. In addition, simultaneous communication and ranging systems reduce latency by saving time compared to traditional stand-alone communication or ranging systems, and overall system performance is improved by reducing hardware complexity. Moreover, simultaneous ranging and communication also contribute to the robustness of the system. By performing both functions simultaneously, noise or interference can be mitigated, and the accuracy of the ranging estimates can be improved.

1.5 Literature Review

As highlighted in Section 1.4.3, the resolution, low data rate, and saturation issues significantly impact system performance. This study places considerable emphasis on addressing the problem of saturation. This section provides a comprehensive literature review of this aspect.

1.5.1 Modulation Methods

A restriction of IS receivers is that the transmission rate is not as fast as that of PDs. Some studies [7, 45, 46] have used high-speed cameras to compensate for the lack of frame rate, with speeds up to 1000 fps. Although high-speed cameras can capture fast motion, it is susceptible to the effects of distance. For high-speed cameras, the size of the transmitter changes depending on the communication distance. To address this challenge, Nishimoto et al. (2012) proposed an overlay coding method in [47, 48]. The method involves distributing short and long-range data across small and large-scale LED arrays and encoding them by superimposing each pattern. The experiments demonstrated that

the overlay coding approach enables error-free communication at distances of up to 70 m. Reference [49] uses multiple exposures to improve further the performance based on the overlay coding scheme in [47]. However, such a design would have limitations according to the number of sizes of LED transmitters.

Pulse width modulation (PWM) is a modulation technique that utilizes multiple voltage levels to increase the data rate [38, 50]. Unlike conventional on-off keying (OOK) modulation that controls the LEDs only in an on/off status, PWM utilizes varying LED brightness levels to represent multiple bits of symbols. Other modulation methods such as pulse-position modulation (PPM) [51], PAM, PSK, QAM, OFDM, and color shift keying (CSK) also effectively increase data rates. However, it is important to note that these modulation methods can be susceptible to LED saturation. Adjusting camera exposure is necessary to capture all LED brightness levels. However, in real-world scenarios, the lighting environment of urban transportation systems can be complex, resulting in LED brightness levels saturating the image and leading to inaccurate range estimations.

In [52], OOK modulation is used. The received 8-bit image shows that the size of the high-level LED is only 7x7 pixels, with 8 pixels saturated at a communication distance of 20 meters. In this case, the luminance can only be estimated using 49 pixels, out of which one-sixth is saturated at a value of 255. In another study [38], Kamegawa et al. propose the application of precoding to PWM voltage levels using a nonlinear LED luminance profile obtained from preliminary experiments. However, since the optimal precoded waveform varies with the communication distance, the performance degrades when using the precoded PWM for moving vehicles. The transmitter can only use the same precoded PWM waveform for all communication distances. Reference [53] employs two cameras with identical configurations for diversity reception, effectively doubling the available

pixel values in the LED area of the received image. However, in this case, the number of PWM levels is limited to three. If we want to increase the data rate by increasing the number of PWM levels, it will become challenging for demodulation due to the limited number of available un-saturated pixels in the LED area. Additionally, saturated pixels result in a nonlinear distribution of the received luminance. Therefore, enhancing the dynamic range is important.

1.5.2 Visible Light Ranging

Much of the ISC literature used stereo-vision-based methods for visible light ranging. In stereo ranging, measurement of the disparity can greatly affect the accuracy of ranging [54]. For more accurate distance measurement, [55] used phase-only correlation (POC) to calculate disparity and sinc function approximation to estimate disparity with subpixel accuracy. Likewise, [52] used equiangular fitting to estimate sub-pixel-level disparity with faster processing than POC. In addition, [56] proposed a neural network (NN)-based localization algorithm, and [57] proposed a scheme to compensate for rolling-shutter effects.

However, there is also literature estimating the range using monocular-vision-based schemes with more than two LEDs. Three LEDs were used in [58] to define the three-dimensional (3D) location of a vehicle in a tunnel. Reference [59] explores the multiple-input-multiple-output scenario, involving more than two LEDs, and adopts the spatial two phase-shift keying modulation technique. In addition, in vehicular applications, solving problems related to vehicle movements is important. Reference [60] used Kalman filters and reduced the random error induced by vehicle movement. Reference [61] introduced a model containing two consecutive image frames to simulate the offset due to

the movement of vehicle vibration. In the two consecutive image frames, two LEDs are placed at the upper left and lower right of the LED array. The image displacements of the upper-left LED in the two consecutive frames were utilized to estimate the offset due to vehicle vibration. It is worth noting that the monocular vision-based scheme in [61] can only be realized with high-speed cameras, since the frame rate should be fast enough to compensate for motion-induced variations.

In addition, machine learning technology has been applied increasingly for visible light ranging to vehicle applications [56, 62, 63]. Reference [56] used a backpropagation NN learning algorithm for vehicle localization. Reference [62] proposed an encoding method for short-range and long-range communication and an artificial NN to predict the positions of the vehicle. In [62], remote communication and accurate localization were performed simultaneously, and the simulation result shows an average range estimation error of 19.8 mm at a communication distance of 30 m.

1.5.3 Simultaneous Visible Light Communication and Ranging

Various studies have explored the potential of simultaneous VLC and ranging, including those referenced in [57,64–67]. In the study conducted in [64], simultaneous communication and ranging were achieved using a spectrum spread code. The system consisted of an LED transmitter and a PD array receiver. Nevertheless, due to PDs' sensitivity to ambient noise, the communication distance is confined to merely 5 m, rendering it inadequate for vehicular applications. Another study [65] employed PDs and LEDs for simultaneous VLC and ranging purposes through Manchester code and measurement of the clock signal's phase shift. However, both [64] and [65] encountered tracking limitations when using PD receivers for outdoor vehicles. The main issue arises from PDs' difficulty in

adapting the optical axis's orientation during high-speed vehicle movement. Augmenting the quantity of PDs in the array is a possible resolution, but building a substantial PD array using existing technologies proves to be formidable.

In contrast, this dissertation adopts ISs as the receivers, which exhibit greater resilience to outdoor ambient noise than PDs. Moreover, IS can easily track vehicles due to their abundant pixels. In addition to their compatibility with various existing car applications, camera-based VLC receivers can also be seamlessly integrated with car recorders, making them highly versatile. Furthermore, the study presented in [66] employed two LED transmitters and an IS to transmit data and measure vehicle-to-vehicle distances. Similarly, in [67], two ISs and one LED transmitter were utilized, resulting in an improved bit error rate compared to using a single IS. Additionally, [57] also explored simultaneous VLC and ranging.

Tables 1.1 and 1.2 summarize the performance in the reference that has been discussed in this dissertation.

Table 1.1 summarizes monocular-ranging performance in the existing literature. It includes references to three studies: Yamazato et al. [61], Takai et al. [68], and Kim et al [58]. These findings highlight the varying results and approaches taken in the studies regarding ranging error (up to 1 m), communication distance ranging from 0 to 60 m, vehicle speed (a constant speed of 30 km/h), and simultaneous communication is achieved only by OCI.

Table 1.2 summarizes the critical information from the references [69] using a high-speed camera and [57] using a rolling shutter camera. Overall, the table provides a concise overview of the ranging error (0.1 m to 1.5 m), communication distance ranging from 0 m to 100 m, vehicle speed ranging from 0 km/h to 100 km/h, and simultaneous commu-

Table 1.1: A summary of monocular ranging performance in the existing literature.

Reference	[61]	[68]	[58]
Ranging Error	0.3 m	N/A	1 m
Communication Distance	30~60 m	around 8 m	0~60 m
Receiver	high-speed camera	OCI	N/A
Vehicle Speed	30 km/h	12.6~14.0 km/h	N/A
Simultaneous Communication	No	Yes	No
Data Rate	N/A	20 Mbps	N/A

Table 1.2: A summary of stereo ranging performance in the existing literature.

Reference	[69]	[57]
Ranging Error	0.5 m	0.1~1.5 m
Communication Distance	20~60 m	0~100 m
Receiver	high-speed camera	rolling shutter camera
Vehicle Speed	N/A	0~100 km/h
Simultaneous Communication	Yes	Yes
Data Rate	500 bps	N/A
Bit Error Rate	Error-free	N/A

nication capabilities described in both a high-speed camera and a rolling shutter camera.

1.6 Purpose of this study

The purpose of the study is to develop a system for simultaneous communication and ranging using ISC and achieve reliable communication and ranging capabilities. To ensure efficient and effective communication and ranging performance, the system should meet specific requirements, including low latency, robust transmission, and high data rate. Additionally, accurate range estimation is crucial for acquiring location information and facilitating vehicle navigation.

As stated in Section 1.5, the investigation of simultaneous communication and ranging in VLC still needs to be improved. One significant challenge in VLC data transmission is the need for higher data rates as the volume of data increases. PWM modulates signals with multiple voltage levels to tackle this challenge, enabling enhanced data rates. This study uses PWM because it is the simplest way to modulate LEDs by controlling luminance and is the least susceptible to LED saturation. However, when dealing with a larger amount of data, it becomes necessary to employ PWM with a greater number of levels to transmit more bits within a single period. Ensuring accurate symbol determination becomes crucial for the successful demodulation of PWM signals. However, using PWM with additional levels introduces complications. Although the transmitted radiance of each level is uniformly distributed, the received LED luminances captured by the cameras exhibit a non-linear distribution due to pixel saturation and the low resolution of the LED. Consequently, accurately measuring luminance becomes less reliable when attempting to differentiate between each level.

This study delves into the novel challenge of addressing LED saturation, which occurs when the luminance of an LED exceeds the dynamic range of the camera. This specific issue has not been explored in the existing literature related to VLC. To tackle this problem, the study proposes two novel approaches:

1. Image processing algorithms, specifically bicubic interpolation techniques, are employed to address the challenges posed by LED saturation in Chapter 3. By analyzing and processing the captured images, the study aims to reconstruct and enhance the information content, resulting in improved accuracy and reliability of the communication and ranging system. It is worth noting that previous works have not explored the use of bicubic interpolations to examine the effects of visible light ranging and LED saturation, making this approach a novel and unexplored contribution in the field.

2. A novel concept is introduced in Chapter 4 by adjusting camera settings to achieve high dynamic range (HDR) in the context of ISC. By optimizing camera parameters, such as the aperture, the system can capture a broader range of luminance levels, encompassing both bright LED signals and the surrounding scene. This HDR capability enables more precise and detailed measurements, significantly enhancing the overall performance of the communication and ranging system. Notably, the HDR concept presented in this study is not found in the existing literature, making it a novel contribution to the field of ISC.

By combining these novel approaches, this study strives to surmount the challenges associated with LED saturation and elevate the capabilities of communication and ranging systems. The ultimate goal is to ensure robust and reliable performance, even in

challenging lighting conditions, thus enabling various applications, including intelligent transportation systems, robotics, and augmented reality.

1.7 Structure of this Dissertation

This chapter provides an overview of the research trends in VLC and ISC based on statistical data gathered from Scopus. It highlights the advantages of utilizing ISs in outdoor environments and delves into the research progress of ISC in the context of ITS. Additionally, it outlines the existing works and the research objective. The remainder of this dissertation is structured as follows:

In Chapter 2, an in-depth explanation of a vehicular VLC and ranging system using two high-speed cameras is presented. Section 2.1 provides an overview of the fundamental architecture and concept of the vehicular VLC and ranging system, detailing its underlying structure. Specifically, Section 2.1.1 focuses on introducing the exposure mechanism of the global-shutter IS, and highlights the receiver structure. Section 2.1.2 delves into the stereo-ranging techniques employed in ISC. Then, Section 2.2 describes the simutaneous method.

In Chapter 3, a stereo camera approach is introduced, enabling simultaneous communication and ranging capabilities. Section 3.1 outlines the system model, which involves capturing images of the same field from different perspectives using two identical cameras. Section 3.2 expresses the utilization of bicubic interpolation for LED saturation. Section 3.3.1 focuses on estimating the range by comparing both images and determining the disparity by applying POC and Section 3.3.2 outlines sinc function approximation. Additionally, in Section 3.4, the proposed system achieves communication by demodu-

lating separate data streams from both cameras, employing a diversity reception method to enhance communication performance. The maximum ratio combining (MRC) technique optimizes the system's performance. It is worth noting that conventional studies typically demodulate data from a single camera, but using two cameras in the proposed system allows for improved communication capabilities.

In Chapter 4, an approach utilizing two high-speed cameras with varying exposure setups as receivers in the ISC system is proposed. Section 4.1 explains the proposed system that applies HDR images to increase the data rate and reduce the symbol error rate (SER). By employing two cameras with varying exposures, one can capture a broader range of brightness information within the LED area. The camera with a higher exposure value is expected to capture finer details of LEDs with lower brightness, whereas the camera with a lower exposure value should capture more details of LEDs with higher brightness. As a result, integrating images of LEDs with both high and low exposure values facilitates more precise PWM demodulation. Section 4.2 describes the HDR demodulation process. The HDR combining technique determines the symbols that correspond to the brightness values of the LEDs. Finally, Section 4.3 presents the results obtained from the field trial experiments for the cases of PWM-8 and PWM-16.

Finally, Chapter 5 presents the summary of this dissertation and the future developments.

Chapter 2

Vehicular Visible Light Communication and Ranging

This chapter presents an in-depth explanation of vehicular VLC and ranging systems. Section 2.1 provides an overview of the fundamental architecture and concept of the stereo-vision-based vehicular VLC and ranging system, detailing its underlying structure. Specifically, Section 2.1.1 introduces the receiver structure and exposure mechanism of global-shutter IS. Section 2.1.2 delves into the ranging techniques employed in ISC, and discusses the range estimation scheme using stereo cameras. Then, Section 2.2 expresses the simultaneous VLC and ranging methods.

2.1 Vehicle-to-everything communications using two high-speed cameras and LEDs

ISC can be applied in various scenarios of ITS, such as vehicle-to-everything (V2X) communication. V2X, including V2V, V2I, and I2V communication, involves interactions between a vehicle and its nearby objects. Many kinds of information can be transmitted in ISC-based V2X communications. The followings are a few application scenarios:

1. ISC enables vehicles to exchange speed and location data, enhancing security and optimizing traffic efficiency. For instance, when a car detects a road obstacle, it can relay this data to the cars behind it, issuing an alert and enabling them to take necessary actions.

2. Transportation facilities can use ISC to transmit information about the traffic environment, such as congestion status or any road dangers, to vehicles to improve security and efficiency. For instance, a traffic signal can convey information to the vehicle about the time that is left before the signal changes, enabling vehicles to modify their speed accordingly.

3. Various infrastructural elements on the roadway can relay data to vehicles concerning traffic patterns, road closures, and weather updates. For instance, road-embedded sensors can sense if the road is icy or snowy and convey this critical information to the vehicle, aiding in safe navigation.

This study is focused on achieving simultaneous ISC and ranging by utilizing two high-speed cameras to capture the LEDs. As illustrated in Figure 2.1, both cameras receive the data concurrently and independently estimate the range. The use of high-speed

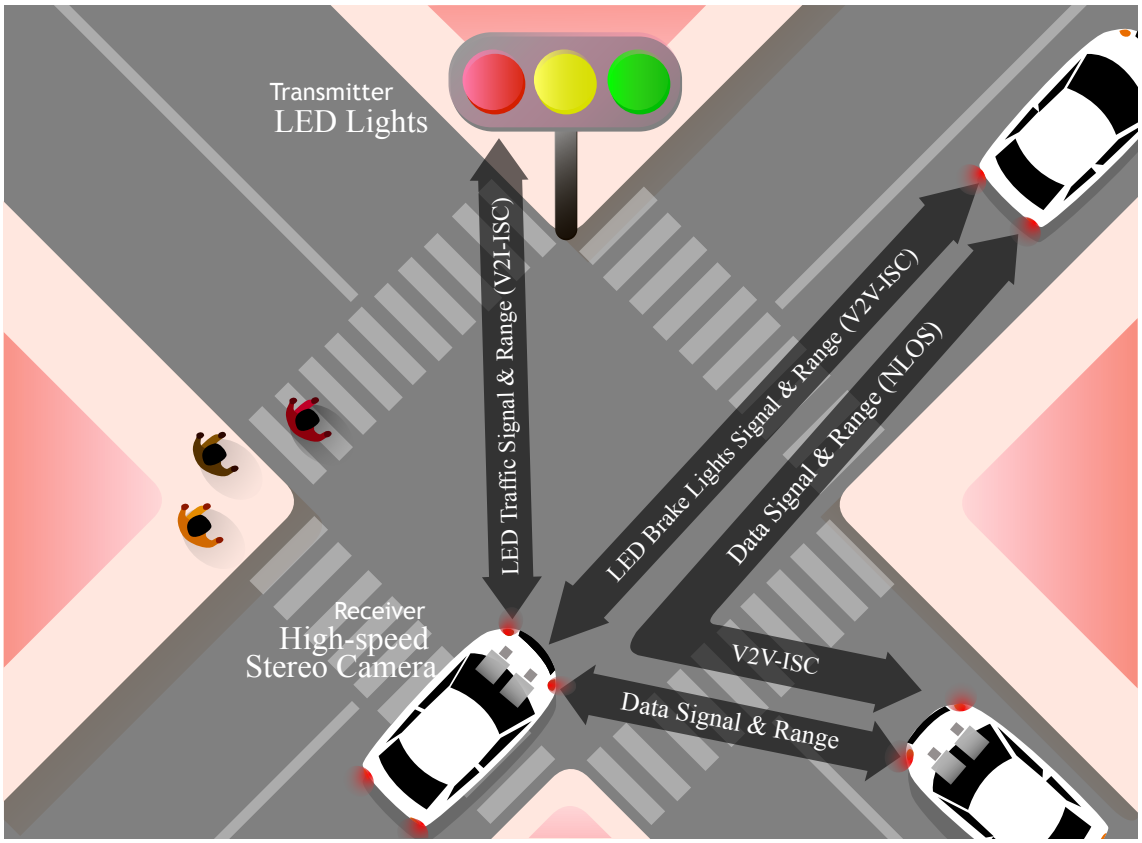


Figure 2.1: The system transmits data signals from light sources such as traffic lights, vehicle headlights, and taillights. Two image sensors are installed on the vehicles to receive data and estimate the range simultaneously. The image sensor captures the visible light signals and converts them into electrical signals. These electrical signals are then decoded using image processing techniques.

cameras ensures fast image capture, contributing to the overall reliability of the system.

ISC can only communicate within the line-of-sight (LOS), i.e., if an obstacle blocks the light, it will interrupt the data transmission. This situation can occur in many applications, such as the case when a neighboring vehicle changes the lane and moves to the front. To enable ISC in non-line-of-sight (NLOS) situations, we can use reflection-based methods [70] or receive data by continuous transmission from the surrounding ISC transmitters. However, on the other hand, the high directivity makes the transmission of VLC more secure and can be used for confidential transmission.

2.1.1 Exposure mechanism of high-speed cameras

This study used high-speed cameras, which have global-shutter ISs.The exposure mechanism of a global-shutter IS is shown in Figure 2.2. In this setup, all rows of the IS expose simultaneously and are then read out line by line. Figure 2.3 illustrates a global-shutter camera capturing an image of a traffic light. Assuming the LEDs blink at the same rate as the readout speed of each row, the image sensor receives an image with all LEDs illuminated. Subsequently, the captured image is transferred to an image-processing computer for further analysis and processing. However, in practical scenarios, the LED and the image sensor are often not synchronous. Therefore, the LED switching speed needs to be equal to or lower than half the camera's frame rate, as dictated by the Nyquist criterion. In this study, the cameras operate at a high frame rate of 1000 fps, while the LEDs are set to blink at a frequency of 500 Hz, which is half the frame rate of the cameras.

The impact of vehicle motion on high-speed camera systems was explored in a study by Kinoshita et al. (2015) [8]. The research employed a pinhole camera model to map world coordinate onto image coordinate and investigated three ISC systems: I2V-ISC,

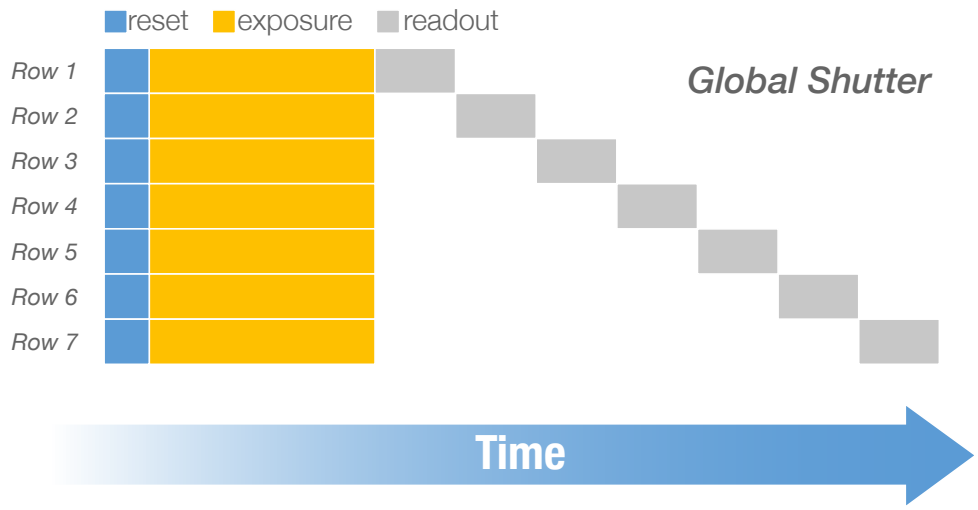


Figure 2.2: The global-shutter exposure scheme is illustrated over time, wherein rows are simultaneously exposed and sequentially read out.

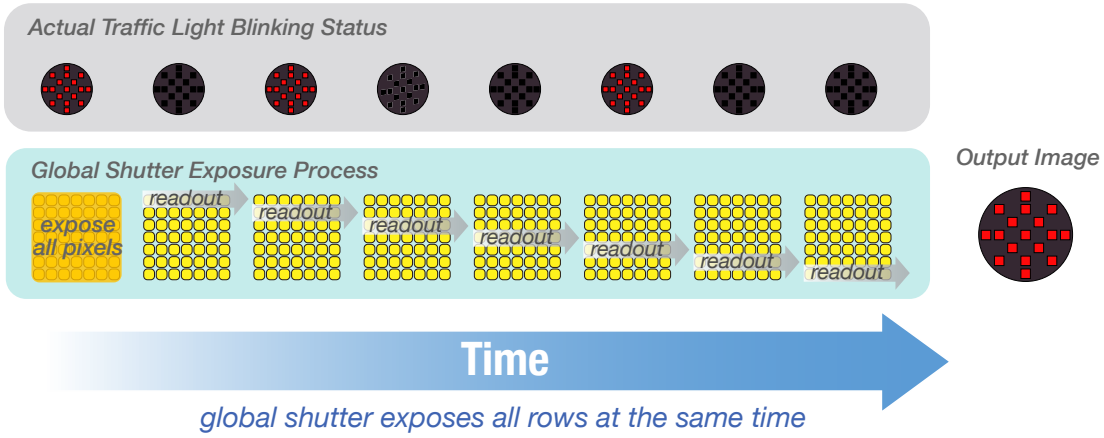


Figure 2.3: When a blinking LED array, as shown above, is captured by a global-shutter image sensor, the output image will exhibit the same as the first LED array pattern. It is assumed that the switching speed of the LED matches the readout rate of the sensor.

V2I-ISC, and V2V-ISC. In the I2V-ISC scenario, the camera traveled together with the vehicle, while the transmitter remained stationary. Conversely, in V2I-ISC, the camera was static, and the transmitter traveled with the vehicle. The study also compared vehicle motion models for I2V-ISC and V2I-ISC and analyzed the effects of camera posture on these models. Additionally, the article discussed how the relative distance between the transmitter and receiver affects the apparent size and position of objects in the captured images [8].

The frame rates exceeding 1000 fps of high-speed cameras result in minimal differences in the transmitter's position between consecutive frames. For instance, if a vehicle travels at 40 km/h, and the camera operates at 1000 fps, the vehicle would move only 0.011 m during each frame, typically reflected within a single pixel on the image plane. This characteristic allows high-speed cameras to excel in vehicle tracking, capturing less offset in the location of adjacent frames compared to rolling shutter cameras.

In this study, the LED detection and tracking method proposed in [71] was employed. In this approach, spatial gradient and temporal gradient played a key role. Spatial gradients involved the computation of horizontal and vertical gradients in the current image frame, particularly in regions where LEDs needed to be recognized, achieved through the Sobel operator. On the other hand, temporal gradients were obtained by assessing the gradients between the current image frame and its preceding and succeeding frames, also utilizing the Sobel operator. LED arrays were identified by analyzing regions with high spatial-gradient and temporal-gradient values.

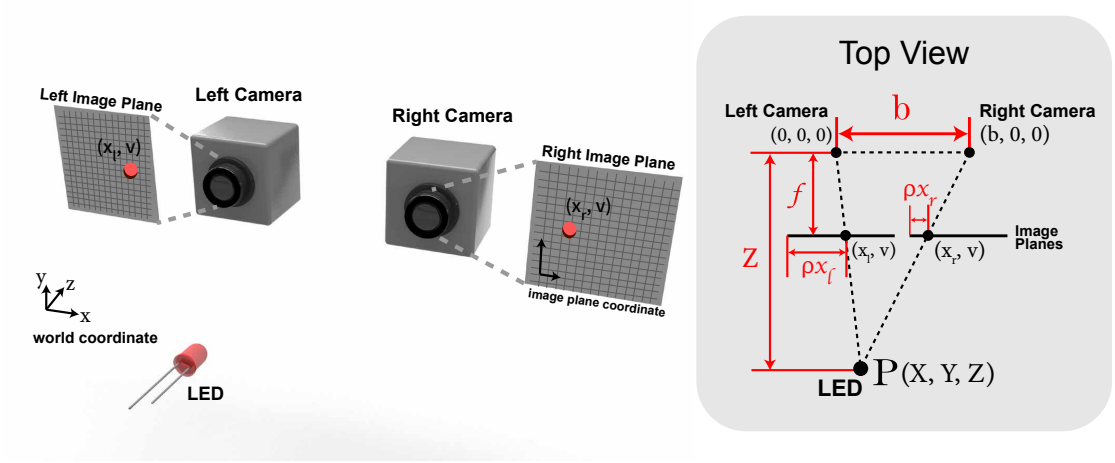


Figure 2.4: A diagram illustrating two calibrated cameras capturing a single LED is presented. The top view is displayed on the right-hand side. In this diagram, b denotes the distance between the left and right cameras, f signifies the camera's focal length, ρ represents the size of one pixel, P represents the LED, and Z indicates the distance that needs to be estimated.

2.1.2 Stereo Vision-Based Range Estimation

This section outlines the fundamentals of stereo vision-based ISC ranging methods. IS can obtain depth information by triangulation. The triangulation method exploits the IS's spatial separation property between different viewpoints to estimate the relative locations of the LEDs in the field. By determining the distance between the LEDs and the camera, range estimation can be realized. The distance between the LED and the camera can be deduced by multiple captures of the same field taken from separate perspectives, particularly the left and right cameras [72–74]. This approach is employed in the stereo vision-based ISC ranging scheme, which employs two camera receivers. The 3D geometry is used to infer the distance between the LED and the camera.

Figure 2.4 illustrates the principle of ranging using two cameras and one LED based on the pinhole camera model. It is assumed that the two cameras are identical and need to be calibrated beforehand. Range estimation using a single LED is feasible in this configuration, where the cameras function as on-board cameras and the LED is mounted on a lead vehicle or traffic light.

Assuming that the two cameras are positioned exactly parallel to each other and facing the LED. The LED is located at point P(X, Y, Z) in the world coordinate system, where the origin point of the world coordinates is at the principal point of the left camera. Left and right camera is at $(0, 0, 0)$ and $(b, 0, 0)$ in the world coordinates, respectively. When both cameras capture the same feature point P of the object, images of point P are received separately on the image planes of the left and right cameras. The coordinates of the feature point can be represented as $\mathrm{P}_{\mathrm{left}}(x_l, v)$ and $\mathrm{P}_{\mathrm{right}}(x_r, v)$ on the image planes of the left and right cameras, respectively, with the origin points of both image planes at the centers of the planes. It is important to note that P, P_{left}, and P_{right} are represented in

different coordinate systems. If we use the properties of similar triangles, we can obtain

$$\begin{cases} x_l = f\dfrac{X}{Z} \\ x_r = f\dfrac{(b-X)}{Z} \\ v = f\dfrac{Y}{Z} \end{cases}. \tag{2.1}$$

where f denotes the focal length of the cameras, b is the distance between the two cameras, and Z denotes the estimated range between the object and the cameras. The LED is projected into (x_l, v) and (x_r, v) on the image planes of the left and right cameras, respectively. If we denote the disparity as d and the parameter to convert disparity to the actual size of a single pixel as ρ, then d can be represented as $d = \rho(x_l - x_r)$. We can then calculate the range using the following equation:

$$Z = f\frac{b}{d} = f\frac{b}{\rho(x_l - x_r)}. \tag{2.2}$$

Therefore, by identifying the corresponding feature point in the image planes of the left and right views, the precise position of the LED can be determined. It is important to note that the aforementioned calculation relies on two feature points, each represented by a single LED. In practical scenarios, the saturation of LED luminance can cause light diffusion, posing a challenge for the receiver to accurately estimate the position of the single feature point.

2.2 Simultaneous Ranging and Communication

There are several approaches available for achieving simultaneous ranging and communication. One method involves employing multiple IS receivers to receive the data signal

and estimate the range using stereo vision techniques [63, 69]. The system can gather data from different perspectives and perform accurate ranging calculations by utilizing multiple receivers.

Another technique involves encoding both data and ranging information in multiple ISC transmitters. These encoded signals are then received and decoded by a single IS receiver [68]. This approach allows for the consolidation of data and ranging information, simplifying the system architecture.

The choice of a specific method depends on the requirements and limitations of the system, as well as the desired trade-off between communication efficiency and ranging accuracy. Each approach has its own advantages and considerations, and the selection should be made based on the specific needs and constraints of the application.

2.3 Summary of Chapter 2

This chapter provides a comprehensive demonstration of vehicular VLC and ranging systems. Section 2.1 introduces the fundamental architecture and concept of the vehicular VLC and ranging system, offering a detailed explanation of its underlying structure. Section 2.1.1 presents the IS receivers utilized in this study: global-shutter IS. Moving on to Section 2.1.2, an in-depth analysis of the ISC stereo-ranging scheme is provided. Finally, in Section 2.2, the simultaneous communication and ranging techniques employed in ISC are explored, discussing their principles and applications in the context of ISC.

Chapter 3

Simultaneous Visible Light Communication and Range Estimation Considering LED Saturation

As mentioned in Chapter 1, the simultaneous integration of ISC and ranging offers performance improvements in both aspects. The main objective of this chapter is to achieve high data rates and accurate ranging while addressing the issue of LED saturation. To tackle LED saturation, a solution is proposed using POC and the bicubic interpolation scheme. For a comprehensive understanding, the fundamental principles, including receiver structure and stereo ranging, are detailed in Chapter 2.

3.1 System Model

Figure 3.1 illustrates the proposed system model, containing both the transmitter and receiver components. This system aims to achieve high data rates through PWM modulation and accomplish accurate ranging simultaneously. As previously discussed in Chapter 1, the main challenge faced by this system is LED saturation. The contributions of this system lie in the introduction of bicubic interpolation and POC methods. The bicubic interpolation helps address the issue of LED saturation by increasing the sampling points.

3.1.1 Transmitter

The header, training, and data sequences constitute the input VLC data. In the transmitter, PWM modulates the data, which is achieved by controlling the duty cycle of the entered electrical signal to transmit M-level symbols. Here, we can denote PWM that has M-level symbols by PWM-M. In Figure 3.2 (a), the symbols 1-4 signals of the PWM-4 are illustrated, with V_{LED} representing the LED forward voltage and T representing the optical clock duration time for the highest-level PWM symbol. Let D be the pulse duration time, and it can take on values in the set $0, T/(M-1), 2T/(M-1), \ldots, T$, where M is greater than 2.

3.1.2 Receiver

The receiver captures the modulated LED optical signal using two synchronized, pre-calibrated cameras [75]. The use of two cameras enables simultaneous communication and ranging, with each camera operating independently. LED position detection is initially performed using a method described in [71]. For the communication function, the

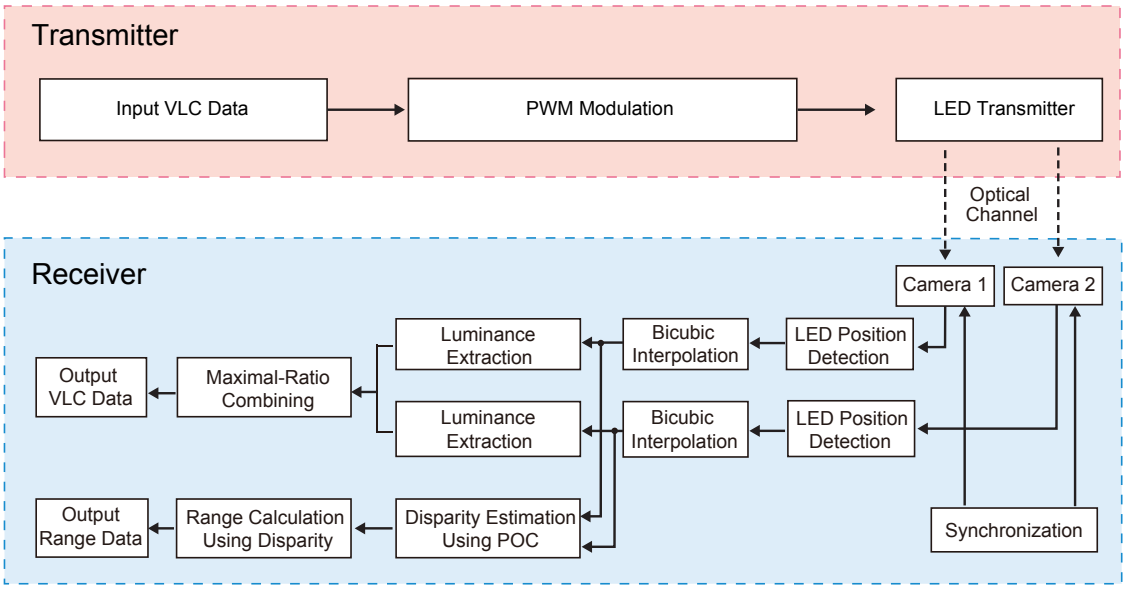
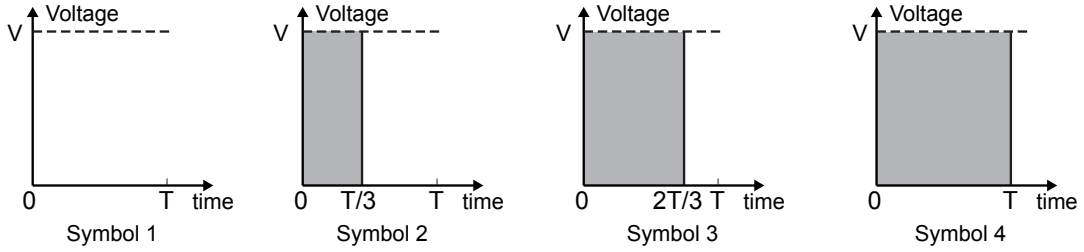
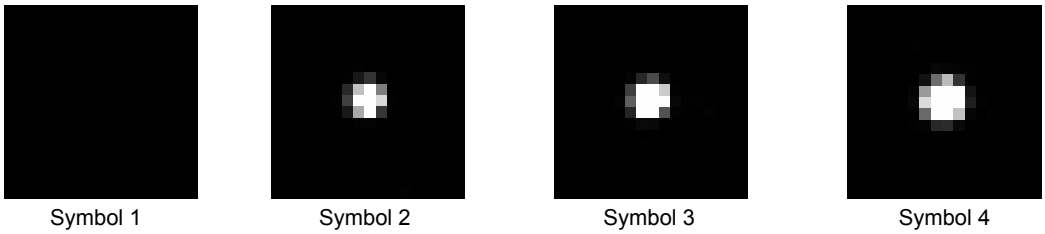


Figure 3.1: Simultaneous visible light communication and range estimation system model. The upper part of the receiver is for the VLC data reception, and the lower part estimate the range.



(a) Transmitted Electrical Signals



(b) Received Images

Figure 3.2: The transmitter side and receiver side for PWM modulation with a modulation level of 4.

VLC data is recovered by extracting luminance [38] and demodulating PWM signals using MRC techniques [53]. Regarding the ranging function, the POC and sinc functions are matched to estimate the disparity. A physical ranging model calculates the distance between the camera and the LED.

Figure 3.2(b) illustrates the LED images with different brightness levels (1-4) of PWM-4 at the communication distance of 60 m. Symbol 1 represents the least saturated LED (completely dark image), while symbol 4 represents the most saturated LED. Diverse LED luminance levels result in varying accuracy in disparity estimation due to differences in LED sizes and diffusion levels between saturated LED and unsaturated LED images. The LED brightness level significantly impacts communication performance and range estimation accuracy. Hence, it is crucial to develop an algorithm that can handle LEDs with different brightness levels.

3.2 Bicubic Interpolation

Following the LED detection and preceding the image processing for both communication and ranging, bicubic interpolation is utilized to address the LED saturation issue. Bicubic interpolation is specifically employed for an individual LED. Figure 3.3 displays the process of capturing an LED and the bicubic interpolation of a single LED, where a point $(x + \Delta_x, y + \Delta_y)$ in the received image corresponds to the point (x_{bic}, y_{bic}) in the interpolation image. The relationship between these two points is expressed as follows:

$$\begin{cases} x_{bic} = \alpha(x + \Delta_x) \\ y_{bic} = \alpha(y + \Delta_y) \end{cases}, \tag{3.1}$$

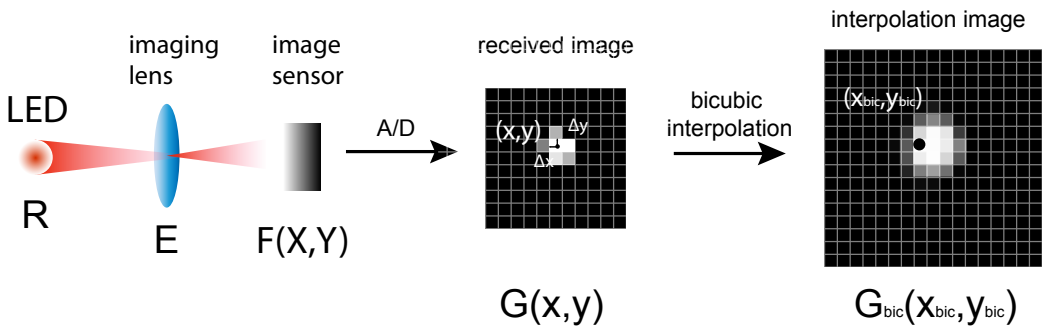


Figure 3.3: The process of receiving an LED and interpolate signal. $G(x, y)$ represents the received image and $G_{bic}(x_{bic}, y_{bic})$ represents the interpolation image. (x, y) and (x_{bic}, y_{bic}) are the corresponding points when $G(x, y)$ is interpolated into $G_{bic}(x_{bic}, y_{bic})$.

where α is the magnification factor, (x, y) and (x_{bic}, y_{bic}) are integer values, Δ_x and Δ_y are noninteger subpixel shifts that are less than one pixel.

The following three steps are used to conduct the bicubic interpolation for a single LED in Figure 3.3:

1. The pixel value of the point $(x_{bic},\ y_{bic})$ in the interpolation image is calculated by determining the coordinate $(x + \Delta_x,\ y + \Delta_y)$ in the received image, which maps the pixel point to the interpolation image. $(x + \Delta_x,\ y + \Delta_y)$ is calculated by: $x = \lfloor x_{bic}/\alpha \rfloor, \Delta x = \{x_{bic}/\alpha\}, y = \lfloor y_{bic}/\alpha \rfloor, \Delta y = \{y_{bic}/\alpha\}$. The operator $\lfloor\rfloor$ represents the floor function, which means to take the largest integer that is less than or equal to the given value. The operator $\{\}$ denotes the fractional part, which represents the decimal part that exceeds the integer part of the given value.

2. Multiply the value of the adjacent 16 pixels of $(x + \Delta_x, y + \Delta_y)$ in the received image with the bicubic interpolation kernel.

3. The value corresponding to the point $(x_{bic},\ y_{bic})$ in the interpolated image is given by the sum of the weighted multiplications of the adjacent 16 pixels of $(x + \Delta_x,\ y + \Delta_y)$ and bicubic interpolation kernel in the received image.

The interpolation image $G_{bic}(x_{bic}, y_{bic})$ is given by

$$G_{bic}(x_{bic}, y_{bic}) = \sum_{h=-1}^{2} \sum_{v=-1}^{2} G(x + h, y + v) S(h - \Delta_x) S(v - \Delta_y), \tag{3.2}$$

where $G(x, y)$ is the input LED image, and h and v represent the positions on the horizontal and vertical axis, respectively.

The bicubic interpolation kernel $S(u)$ is given by

$$S(u) = \begin{cases} 1 - 2.5u^2 + 1.5|u|^3, 0 \leq |u| \leq 1 \\ 2 - 4|u| + 2.5u^2 - 0.5|u|^3, 1 < |u| \leq 2 \\ 0, \text{otherwise} \end{cases} \tag{3.3}$$

where u denotes the position difference between the interpolation point and the current pixel [76].

3.3 Ranging

3.3.1 Phase-only Correlation

POC is employed to determine the disparity, i.e., (x_l-x_r) in Figure 2.4 (Section 2.1.2), for estimating the distance between cameras and LEDs. POC is a scheme that calculates the phase correlation to assess the relative translation offset between the images captured by the left and right cameras.

Suppose G_{left} and G_{right} are the interpolated signals of the images captured by the left and right cameras, respectively. The relationship between both signals can be represented as

$$G_{right}(x_{bic}, y_{bic}) \cong G_{left}(x_{bic} + \Delta(x_{bic}), y_{bic} + \Delta(y_{bic})), \tag{3.4}$$

where $\Delta(x_{bic})$ denotes the image displacement on the x-axis, i.e., the disparity, and $\Delta(y_{bic})$ denotes the image displacement on the y-axis. In this scenario, the ideal case is when the two images perfectly match, and the displacement between them represents the disparity. As calibrated two cameras are used in this study, $\Delta(y_{bic}) \cong 0$.

The cross-power spectrum of the left and right images can be defined as

$$C\left(x_{bic}, y_{bic}\right) = \frac{G_{left} G^{*}_{right}}{|G_{left} G^{*}_{right}|} = e^{j\theta(k_1, k_2)}. \tag{3.5}$$

The POC function $c(x_{bic}, y_{bic})$, inverse discrete Fourier transform of the cross-power spectrum of $C(x_{bic}, y_{bic})$, is given by a Kronecker's delta function δ_K:

$$c(x_{bic}, y_{bic}) \cong \delta_K(x_{bic} + \Delta(x_{bic}), y_{bic}), \tag{3.6}$$

where the peak of δ_K is at $(\Delta(x_{bic}), 0)$ [77]. Thus, we can determine the disparity $\Delta(x_{bic})$ by searching for the peak of $c(x_{bic}, y_{bic})$:

$$(\Delta(x_{bic}), 0) \cong \underset{(x_{bic}, y_{bic})}{\arg\max}(c) \tag{3.7}$$

Figure 3.4 shows a 3D example of the POC function taken at a communication distance of 20 m. In this example, the disparity is determined to be 105 pixels with the center of the IDFT at (256, 256).

3.3.2 Sinc Function Matching

Estimating the disparity with subpixel accuracy becomes crucial because the actual LED center does not align precisely with a pixel-level position. To estimate the position displacement of two images at the subpixel rather than pixel level, we can rewrite (3.6) as follows:

$$c(n_1, n_2) = \frac{1}{WH} \frac{\sin(\pi(n_1 + \delta_1))}{\sin(\pi(n_1 + \delta_1)/W)} \cdot \frac{\sin(\pi(n_2 + \delta_2))}{\sin(\pi(n_2 + \delta_2)/H)}, \tag{3.8}$$

where W and H denote the width and height of the image, respectively, and δ_1 and δ_2 denote the subpixel image displacement on the x and y-axis, respectively.

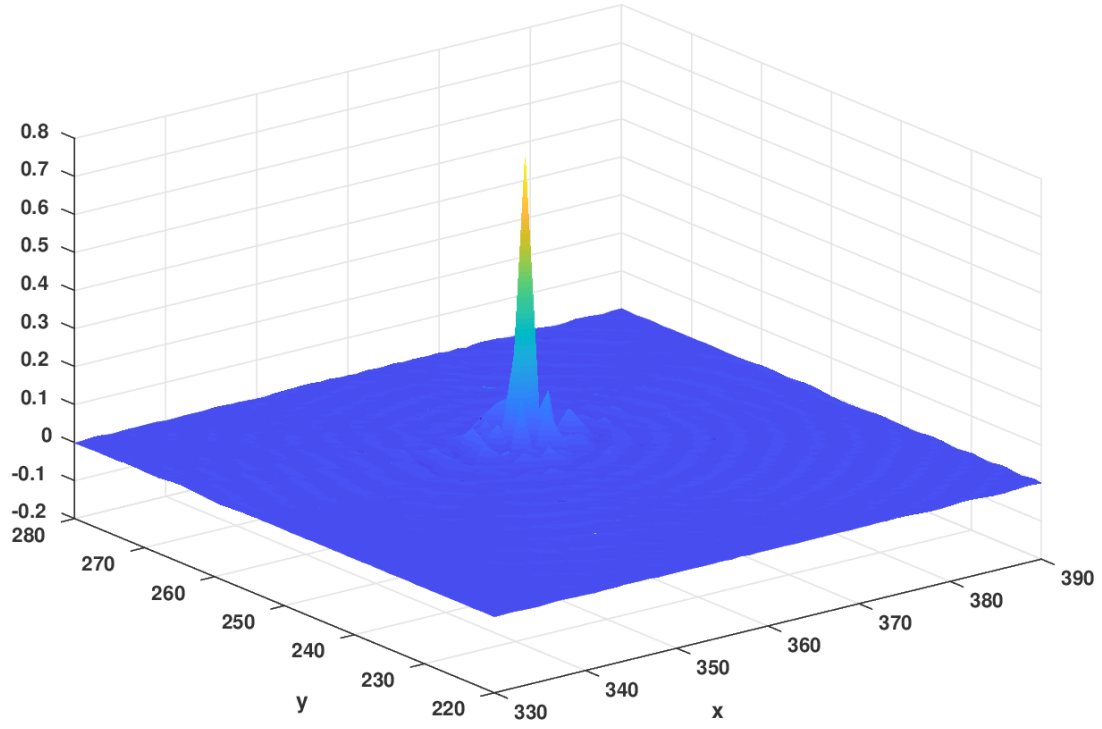


Figure 3.4: The POC function taken at a distance of 20 m, where the correlation peak is denoted by the color yellow. x and y are the horizontal and vertical axis of the image, respectively. z is the correlation value.

(3.8) is a very close approximation to the sinc function. Thus, an approximation of (3.8) can be given by

$$c(x_{bic}, y_{bic}) \cong \text{sinc}(x_{bic} + \gamma(x_{bic}))\text{sinc}(y_{bic}), \tag{3.9}$$

where $\text{sinc}(x)$ denotes $\sin(\pi x)/\pi x$, $\gamma(x_{bic})$ denote the image displacement in the subpixel level on the x-axis [78, 79].

A minimum mean square error estimation technique was applied to achieve synchronization between the POC and the sinc function. This approach allowed for the determination of the peak location of the sinc function, enabling subpixel resolution and accurate range calculation. Increasing the sampling rate, which can be achieved through image interpolation, plays a crucial role in reducing the matching error of the sinc function. This is because a higher sampling rate provides more sampling points, resulting in a decreased proportion of saturated values in the discrete LED image signals. During the experiment, it was observed that the size of a single LED was approximately 6×6 pixels at a distance of 20 m, and around 3×3 pixels at a distance of 60 m. Different levels of LED luminance significantly impact the accuracy of disparity estimation. This is primarily due to variations in the size of the LEDs and their diffusion degrees between saturated and unsaturated LED images. Hence, developing an algorithm that can effectively handle LEDs at different luminance levels is essential to ensure accurate ranging.

Examples of measuring the image registration in one dimension for low and high luminance cases are shown in Figure 3.5. The signals from the left and right cameras represent the captured LED signals based on the light diffusion model [9] and the POC functions represent the image registration result. In the high luminance case, if the value of LED signal is over 255, LED saturation occurs. We can see in the graph that the subpixel estimation differ between the two cases.

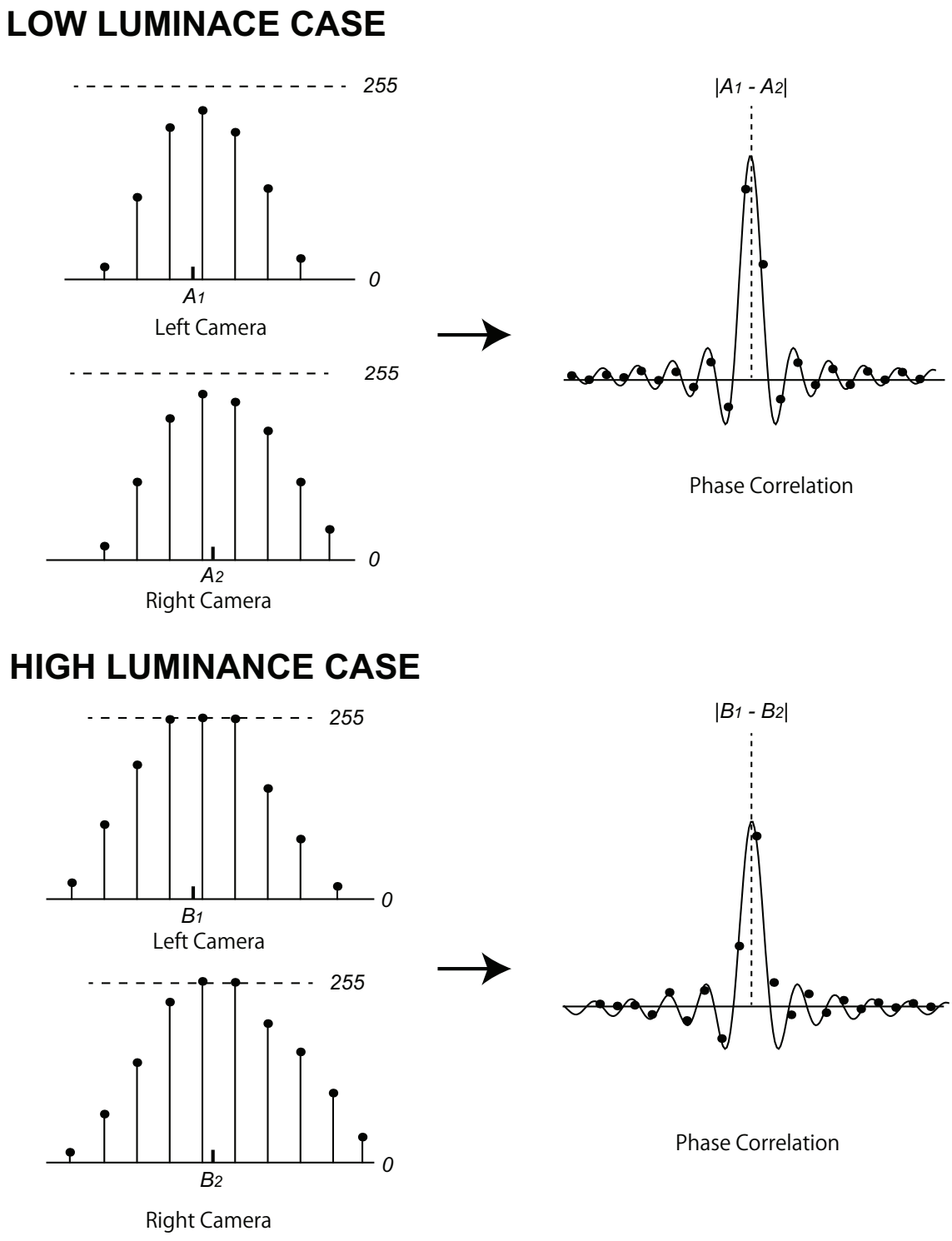


Figure 3.5: Phase correlation of different luminance level LEDs. Using sinc functions to match image pixels of low and high luminance LED has different characteristics.

3.4 Communication

3.4.1 LED Luminance Extraction

The luminance of each LED is measured by averaging the pixel values in the detected LED region [38]. The luminance of an individual LED is denoted by L_i where i is the spatial index of the LED in the LED array. The region of interest (ROI) for L_i, i.e., the projected area of the individual LED in the received image, is represented by a rectangular region obtained from LED detection. This study denotes the coordinates of the top-left corner and the bottom-right corner in the ROI of L_i by (x_i^t, y_i^t) and (x_i^b, y_i^b), respectively. The luminance L_i can be expressed by

$$L_i = \frac{\sum_{x=x_i^t}^{x_i^b} \sum_{y=y_i^t}^{y_i^b} I(x,y)}{(x_i^b - x_i^t + 1)(y_i^b - y_i^t + 1)} \tag{3.10}$$

where i ranges over the number of LEDs and $I(x, y)$ is the pixel value at the coordinate (x, y) in the received LED image.

3.4.2 Maximal ratio combining (MRC)

In the proposed method, MRC performs PWM modulation using images from the left and right cameras. The training sequence calculates the channel factor for each symbol and each camera, thus modeling the channel parameters and inferring the symbols.

The channel factors of the left and right cameras, $w_l^{s'}$ and $w_r^{s'}$, are given by

$$\begin{cases} w_l^{s'} = \frac{1}{s' \cdot N_{s'}} \sum_{n=1}^{N_{s'}} L_l^{s'}[n] \\ \\ w_r^{s'} = \frac{1}{s' \cdot N_{s'}} \sum_{n=1}^{N_{s'}} L_r^{s'}[n] \end{cases}, \tag{3.11}$$

where s' is the transmitted symbol in the training sequence, $N_{s'}$ is the data length of training sequence for symbol s', $w_l^{s'}$ and $w_r^{s'}$ are the channel factors of symbol s' for the left and right cameras, respectively, $L_l^{s'}[n]$ and $L_r^{s'}[n]$ are the n-th received luminance values of symbol s' in training sequence for the left and right cameras, respectively.

The estimated symbol $\hat{s}$ in the data sequence is given by

$$\hat{s} = \underset{s\in[1,2,3,\ldots,M]}{\arg\min} \left((L_l - w_l^s \cdot s)^2 + (L_r - w_r^s \cdot s)^2\right) \tag{3.12}$$

where L_l and L_r are the received luminance values in the data sequence of the left and right cameras, respectively, s is a transmission symbol in the data sequence, and M is the symbol level.

3.5 Experiment

The system's performance was evaluated through a series of field trial experiments. The experimental setup and parameters are depicted in Figure 3.6 and Table 3.1, respectively. A single LED is used to transmit data in the experiment. The ranging performance evaluation was based on the distance versus the average absolute estimation error, with measurements taken at 5 m intervals. Each measurement consisted of analyzing five thousand frames at each communication distance. The subpixel estimation level was set at 0.01 pixels, utilizing a magnification factor of 3.

Figure 3.7 presents the range estimation results. Figure 3.7 illustrates the communication distance versus average absolute estimation error. Applying bicubic interpolation has significantly enhanced the range performance in the range of 55 to 60 m. In [34], ranging was performed using a single camera and a specific LED ranging pattern. In contrast, this study utilized two cameras, resulting in a more robust system that does not require a

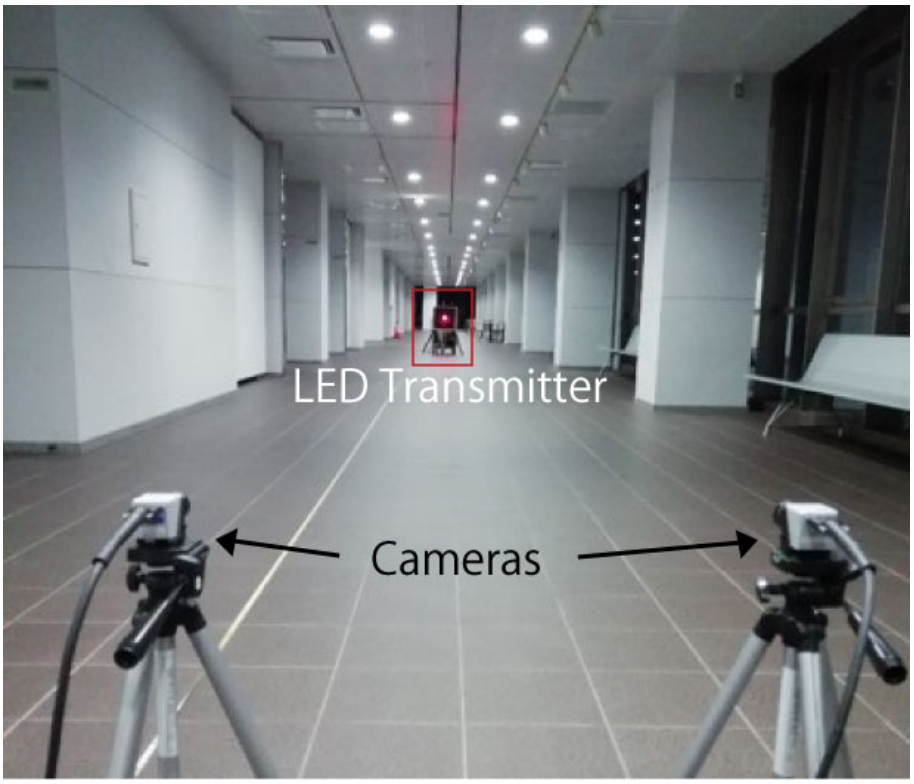


Figure 3.6: Photograph showcasing the experimental setup. The LED transmitter is highlighted within a red rectangle.

Table 3.1: Experiment parameters

Transmitter	
Model	SunLED XZM2ACR55W-3
Frequency	500 Hz
Modulation Method	OOK and PWM8
Receiver	
Model	Photron IDP-Express R2000-F
Exposure Time	1 ms
Frame rate	1000 fps
Pixel size	10 μ m
Resolution	512×512 pixels
Focal length	35 mm
Distance between two cameras	0.6 m
Communication distance	20-60 m, every 5 m

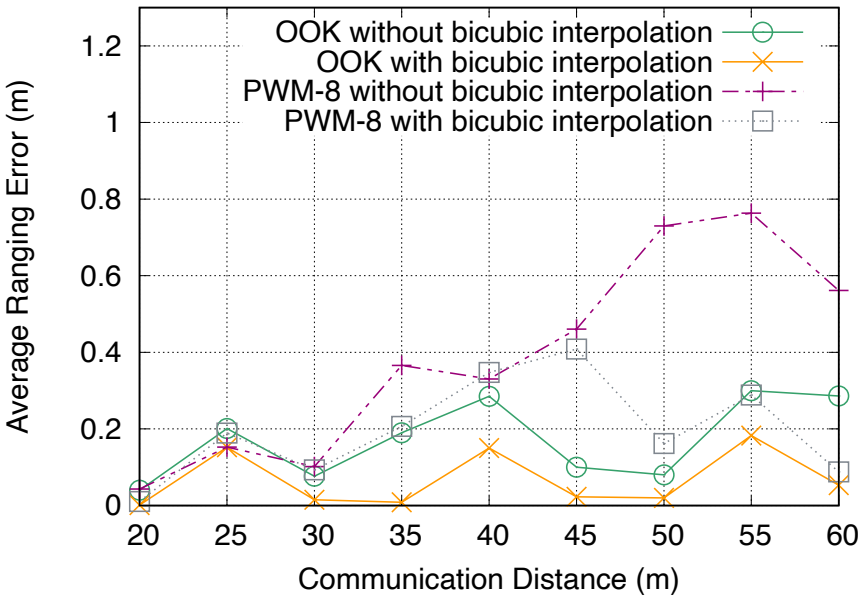


Figure 3.7: Average Range Estimation Error

specific ranging pattern. The estimation error was below 0.5 m within a communication distance of 60 m in both [34] and the current research. With bicubic interpolation, the average absolute range estimation error was below 0.5 m within 60 m, whereas the conventional POC method exhibited an estimation accuracy of approximately 0.8 m. This improvement represents a reduction of approximately 0.3 m in error. Moreover, an accuracy of 0.5 m is sufficient for ITS applications. For instance, when applying the proposed ranging method to the Shibuya crossing, where the distance from the traffic lights to the stop lines is approximately 50 m, the ranging error remains below 0.3 m. It is worth noting that the minimum distance from the stop line to the crosswalk is 1.5 m, as measured by Google Maps. Thus, if the traffic lights transmit ”stop” information to vehicles, an accuracy of 0.3 m is adequate to prevent collisions between vehicles and pedestrians.

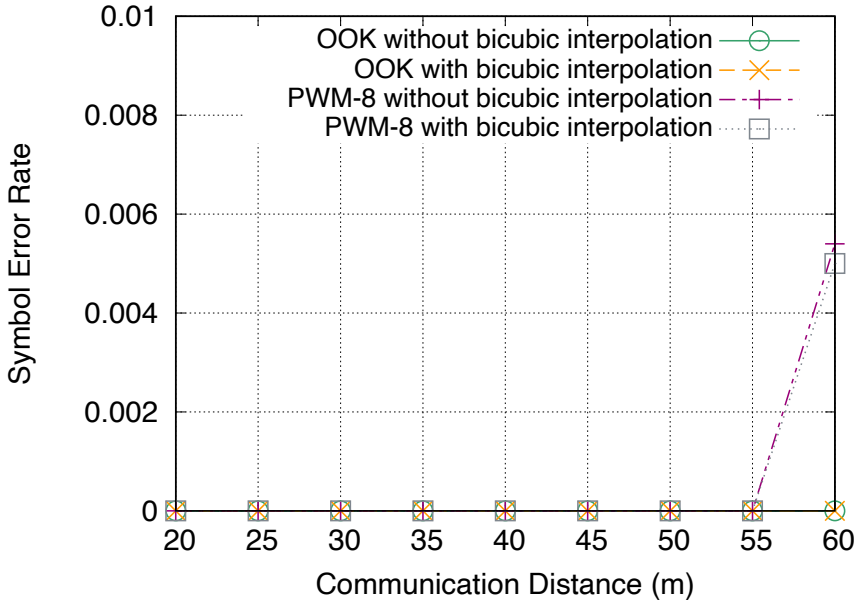


Figure 3.8: The symbol error rate is measured for OOK and PWM-8.

Figure 3.8 displays the SER for PWM-8 modulation compared with OOK. The SER curves for both OOK and PWM-8 indicate error-free performance within a communication distance of 55 m. At a distance of 60 m, PWM-8 achieves an SER of less than 6×10^{-3}, with the interpolation-based approach demonstrating superior performance. This study focuses on enhancing ranging accuracy, while ensuring that the communication performance remains unaffected. The results indicate that applying bicubic interpolation does not degrade the communication performance. Moreover, the data rate of PWM-8 is three times higher than that of OOK, with both achieving error-free performance up to 55 m. This suggests that PWM can significantly increase the data rate while maintaining the same transmitter frequency as OOK. It is important to note that the transmitter frequency

for OOK corresponds to the optical clock rate for an OOK symbol, while the transmitter frequency for PWM corresponds to the optical clock rate for the PWM symbol with the highest luminance level. Assuming the pulse duration time for an OOK symbol and a PWM symbol with maximum brightness is both T, the transmitter frequencies are equal to $1/T$ for both OOK and PWM. Furthermore, a noteworthy observation from our experimental results is that the error-free communication distance of 55 m is suitable for ITS-VLC applications. For instance, in the well-known Shibuya crossing, one of Japan's largest intersections, the maximum distance from the traffic lights to the stop lines is approximately 50 m, as determined by Google Maps.

3.6 Summary of Chapter 3

In this chapter, a simultaneous VLC and ranging system was introduced using PWM modulation, along with a proposed algorithm for disparity estimation in ITS-VLC ranging to address the issue of LED saturation. It was observed that the luminance levels of LEDs substantially impact the accuracy of disparity estimation in VLC stereo-ranging systems. To overcome this, the bicubic interpolation method was employed to mitigate LED saturation, resulting in improved ranging performance and subpixel accuracy in range estimation. Experimental results revealed that the estimation error using the PWM modulation method was less than 0.5 m within a communication distance of 60 m, surpassing the performance of the conventional POC method without bicubic interpolation. Additionally, the communication performance remained error-free up to 55 m.

These findings highlight the potential to achieve high-ranging accuracy in ITS-VLC systems using stereo cameras. Furthermore, the utilization of stereo cameras enables

simultaneous VLC data transmission and range estimations, leading to a significant enhancement in the overall performance of an ITS-VLC system.

Chapter 4

Visible Light Communication Using High Dynamic Range Images

As mentioned in Chapter 2. This chapter addresses the pixel saturation problem by employing a technique to achieve HDR using two high-speed image sensors with different exposure settings. Section 4.1 provides an overview of the HDR approach, while Section 4.2 details the proposed method for signal demodulation using HDR combining. Furthermore, Section 4.3 presents the results of an experiment, demonstrating that a throughput of over 87.5 kbps can be achieved within a communication distance of 50 m under PWM-8. Some basic concepts, including the receiver structure and stereo ranging, are described in Chapter 2.

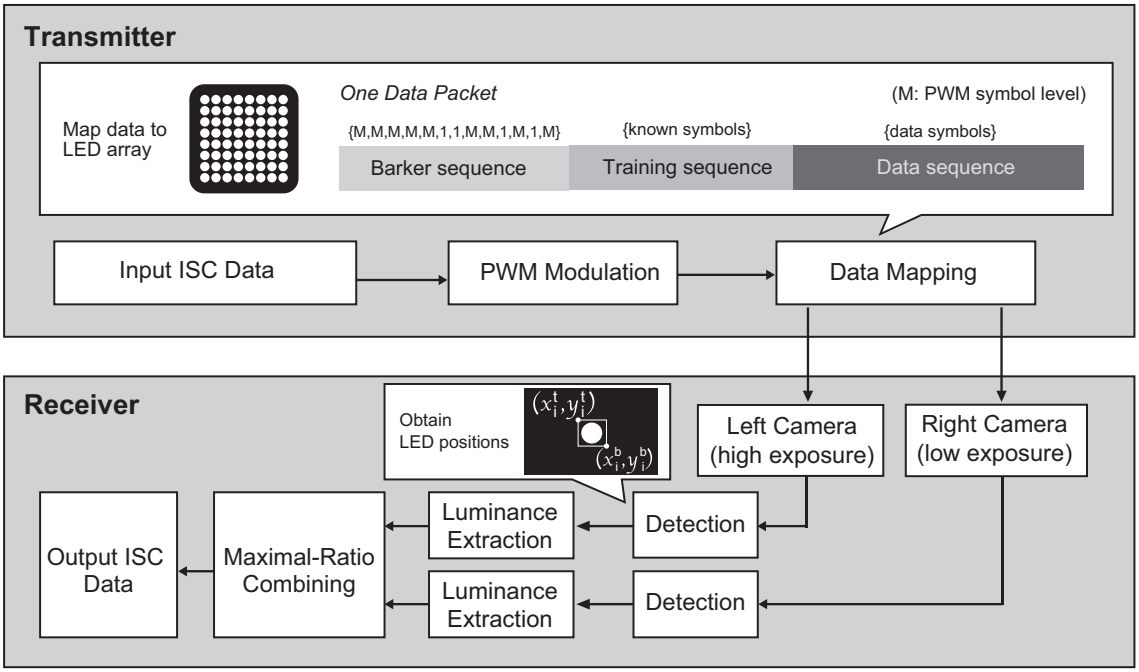


Figure 4.1: System Model for HDR combining.

4.1 Acquiring LEDs using High Dynamic Range (HDR) Settings

The system model of the proposed HDR combining scheme in this chapter is illustrated in Figure 4.1.

As shown in Figure 4.1, first the PWM modulates the information and adds the Barker and training sequences. The sequences are then sequentially assigned to each LED in the transmitter. Each LED will transmit a different signal. In Figure 4.1, M is the number of PWM levels, the Barker sequence consists of thirteen symbols, and the training sequence is a known symbol sequence, the same as in Chapter 3.

At the receiver, the system utilizes two cameras with identical specifications but different exposure settings to capture LEDs. The left camera has a relatively high exposure value, and the right camera has a relatively low exposure value. To adjust the exposure, the apertures of the cameras are changed. In conventional HDR approaches, exposure time is typically adjusted to obtain different exposures [80]. However, in this study, varying the exposure time is not feasible because the cameras need to capture PWM symbols with an exposure time that is half the LED clock time [38]. Therefore, this study adjusts the camera apertures to achieve different exposures. As one of the cameras has a higher exposure and the other has a lower exposure value, the camera with the higher exposure value should be able to capture finer detail of the less bright LED. Conversely, the camera with the lower exposure should capture more detail from the brighter LED. Therefore, combining images of both lower and brighter LEDs can help with more accurate PWM demodulation.

After camera capture, the LED array is initially detected by the method described

in [71] in the subsequent detection process. Subsequently, each LED's position is determined using the contour detection method outlined in [81]. Following the detection process, LED luminance extraction is performed for each LED. Finally, HDR combining is employed to recover the ISC data.

4.2 High Dynamic Range Combining

The simplified channel model for HDR combining is shown in Figure 4.2. It is assumed that the aperture radius for the left and right camera is D_l and D_r, and the channel factors of symbol s of the left and right cameras are h_l^s and h_r^s, respectively.

The received LED luminance is given by

$$\begin{cases} L_l(D_l) = (s \cdot h_l^s)(\pi(D_l)^2) \\ L_r(D_r) = (s \cdot h_r^s)(\pi(D_r)^2) \end{cases} \tag{4.1}$$

Then, if we can estimate the channel factors, we can estimate the symbol. The channel factors for each symbol and each camera are estimated using a training sequence. The training sequence is transmitted before the data sequence using known symbols. $L_l^{s'}[n]$ and $L_r^{s'}[n]$ are calculated using the method in Section 3.4.1. The channel factors of symbol s' of the left and right cameras are denoted by $h_l^{s'}$ and $h_r^{s'}$, respectively. The channel factors are given by

$$\begin{cases} w_l^{s'} = \frac{1}{N_{s'}} \sum_{n=1}^{N_{s'}} \frac{L_l^{s'}[n]}{s'} \\ \\ h_r^{s'} = \frac{1}{N_{s'}} \sum_{n=1}^{N_{s'}} \frac{L_r^{s'}[n]}{s'} \end{cases} \tag{4.2}$$

where s' is a symbol in the training sequence and ranges over the number of levels, $N_{s'}$ is the sequence length for symbol s' in the training sequence, $L_l^{s'}[n]$ and $L_r^{s'}[n]$ are the

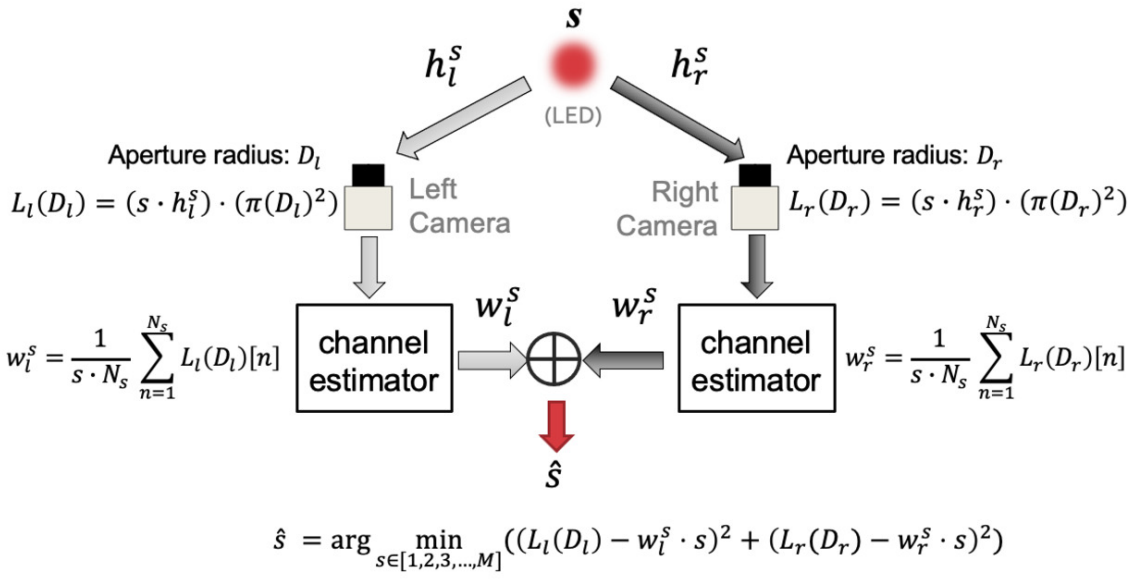


Figure 4.2: A simplified channel model for HDR combining.

extracted luminance values of symbol s' in the training sequence with the time index n for left and right camera, respectively.

Then, the symbol s can be estimated by the following formula:

$$\hat{s} = \underset{s\in[1,2,3,...,M]}{\arg\min} ((L_l(D_l) - w_l^s \cdot s)^2 + (L_r(D_r) - w_r^s \cdot s)^2), \tag{4.3}$$

Please note that the calculation process in this section is the same as described in Section 3.4.2. The difference lies in the aperture parameters used for the HDR combining method.

4.3 Experiment

4.3.1 Experiment setup

The experiments were conducted under indoor daylight conditions. Figure 4.3 visually represents the experimental setup. The LED array transmitter consisted of an 8x8 arrangement of LEDs, separating 6 cm between adjacent LEDs. Each LED sends different data. Two cameras were positioned directly in front of the LED array transmitter to ensure comprehensive coverage, capturing all the LEDs.

During the experiment, the LED array transmitter was moved to different distances, and the cameras captured the transmitted signal to observe the system's performance at varying communication distances. The positions of the cameras remained fixed throughout the experiment to maintain consistent relative positioning. Measurements were collected at 10-m intervals, covering a range from 10 m to 50 m.

Various aperture settings were utilized for the left and right cameras to explore their influence on communication performance. First, the symbol's luminance distributions

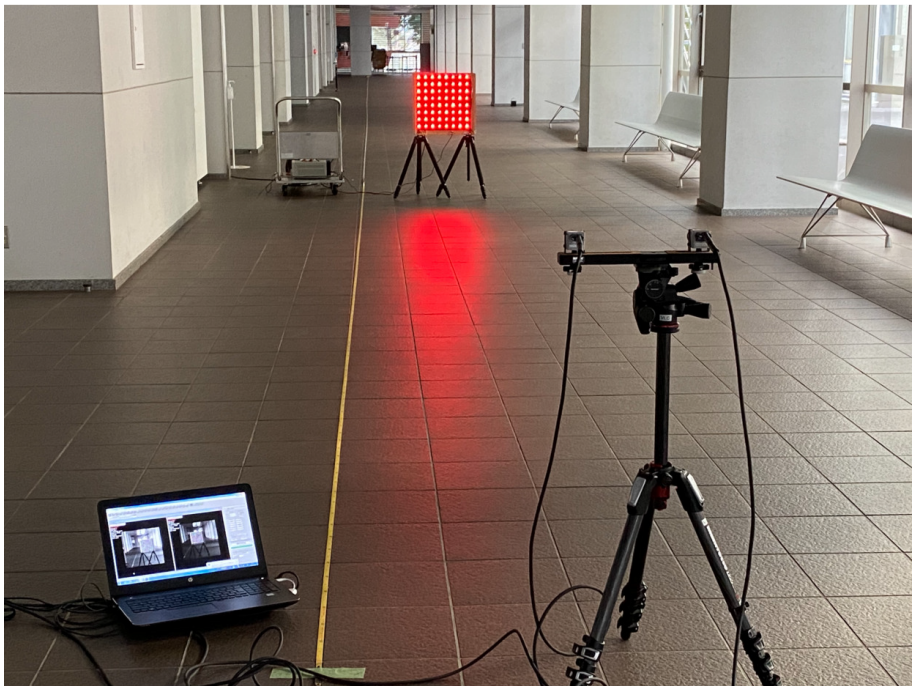


Figure 4.3: Experiment View.

were investigated in Section 4.3.2 for the cases of f/2, f/4, f/8, and f/16. Then, the SER is measured in Section 4.3.3 and 4.3.4. Specifically, the aperture of the left camera was f/4 for all the cases, and the aperture configurations for the right camera were f/4, f/8, and f/16. The f/4 aperture is used for comparison with the non-HDR scenario.

Each measurement involved the transmission of a packet consisting of 653 frames, including 41792 symbols. The packet included a 13-frame Barker sequence, a 40-frame training sequence, and a 600-frame data sequence. SER was evaluated using only the data sequence. Further details regarding the experimental parameters can be found in Table 4.1.

Table 4.1: Experiment parameters

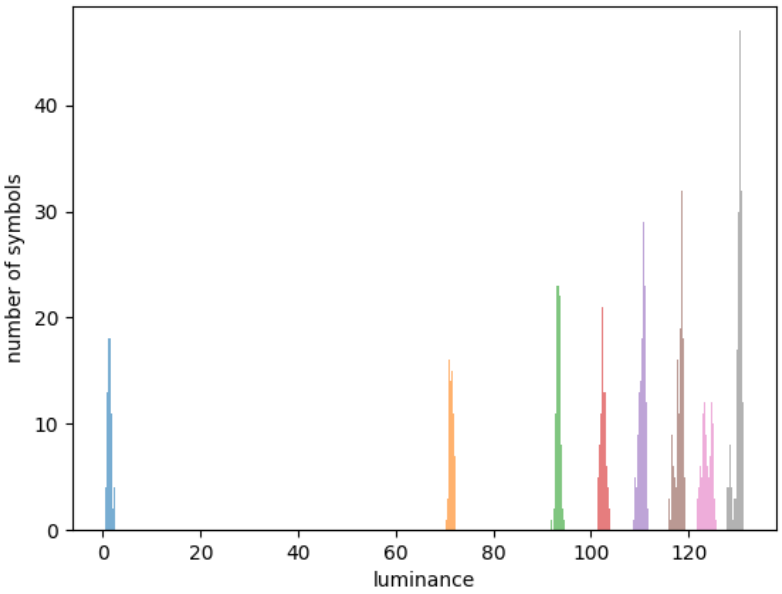
Transmitter	
LED Model	SunLED XZM2ACR55W-3
Number of LEDs	8×8
Blinking Rate	500 Hz
Modulation Method	PWM-8 and PWM-16
Receiver	
Camera Model	Photoron IDP-Express R2000-F
Shutter Speed	1 ms
Frame rate	1000 fps
Image Depth	8-bit image
Resolution	512×512 pixels
Focal length	35 mm
Distance between two cameras	0.25 m

4.3.2 Luminance distribution

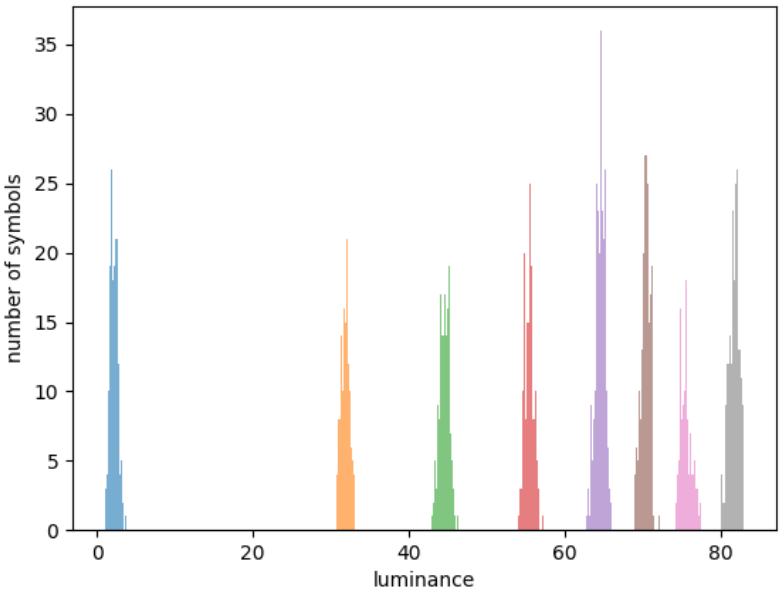
Figure 4.4 shows the luminance distribution for 500 frames at the communication distance of 20 m for PWM-8. The luminance is arranged in an ascending sort. Different colors represent different symbols. The luminance value is measured by pixel value, as explained in Section 4.2. For PWM-8, symbols can be clearly divided and distinguished with apertures of f/2 and f/4, resulting in error-free communication. For apertures of f/8 and f/16, symbols are easy to distinguish with very low SER. However, at aperture f/16, the luminance distribution will be overlapped with the high-luminance value, making it harder to distinguish symbols.

Figure 4.5 show the luminance distribution for 500 frames at the communication distance of 20 m for PWM-16. The luminance is arranged in an ascending sort. For PWM-16, symbols are easily distinguishable with low SER using apertures of f/2 and f/4. However, at apertures of f/8 and f/12, symbols up to 7 can be dealt with, but symbols above 7 become difficult to distinguish. At aperture f/16, distinguishing symbols becomes very hard, resulting in relatively high SER. The luminance distribution with an aperture of f/16 also tends to be linear.

In conclusion, symbols with lower luminance can be easily determined in all cases, while symbols with higher luminance are challenging to determine due to nonlinearity (or LED saturation). The best performance is achieved with apertures of f/2 and f/4, suggesting that using these apertures for the left and right cameras may be a favorable choice. Additionally, the luminance value with an aperture of f/16 exhibits linearity.

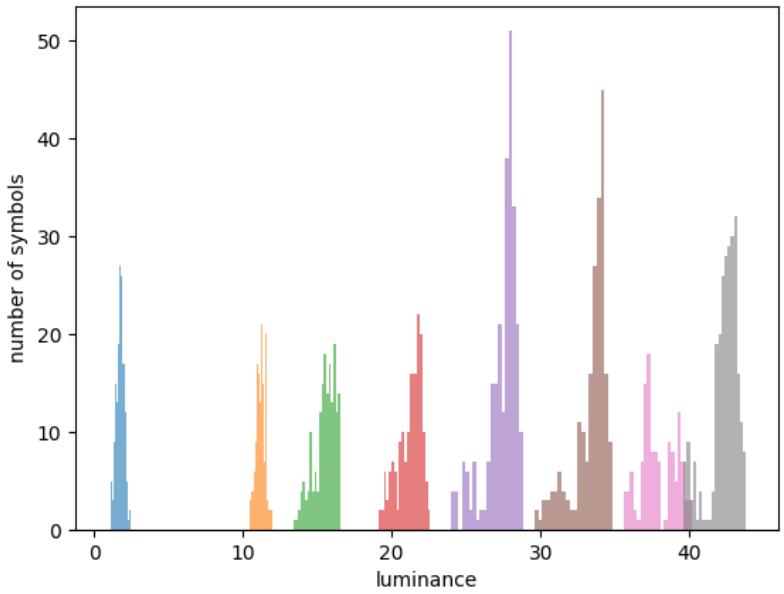


(a)

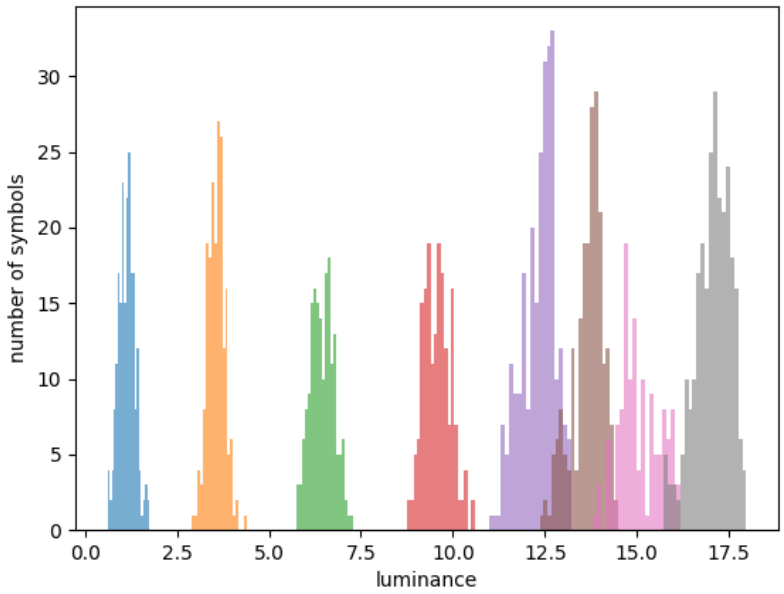


(b)

Figure 4.4: Luminance distributions for PWM-8 at a distance of 20 m. The aperture values are (a) f/2. (b) f/4. (c) f/8. (d) f/16.

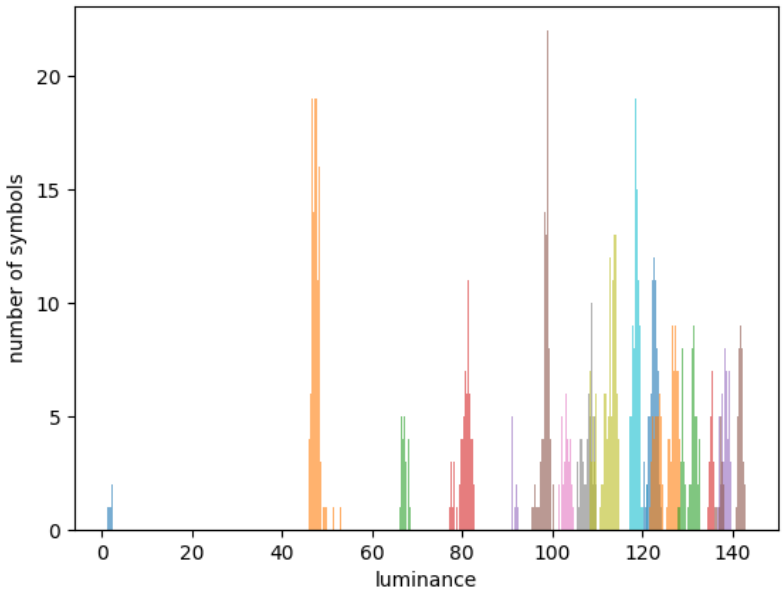


(c)

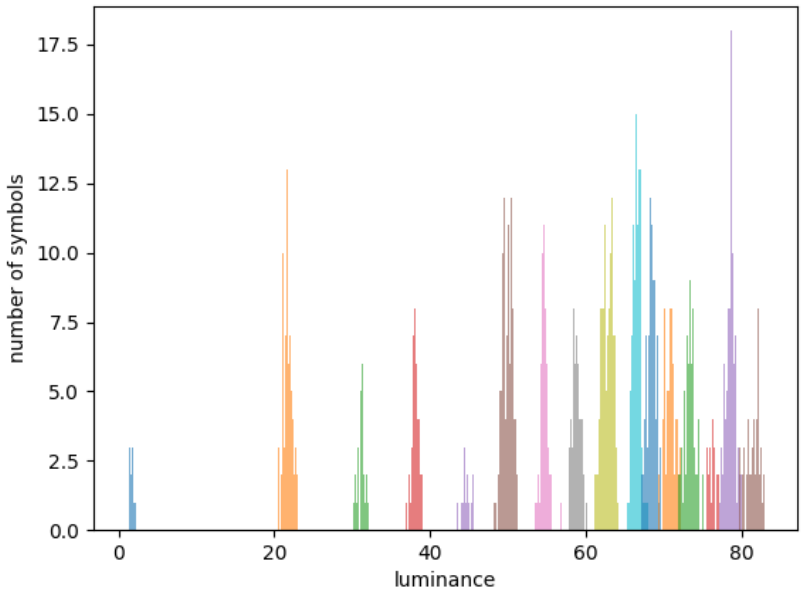


(d)

Figure 4.4: (*continued*) Luminance distributions for PWM-8 at a distance of 20 m. The aperture values are (a) f/2. (b) f/4. (c) f/8. (d) f/16.

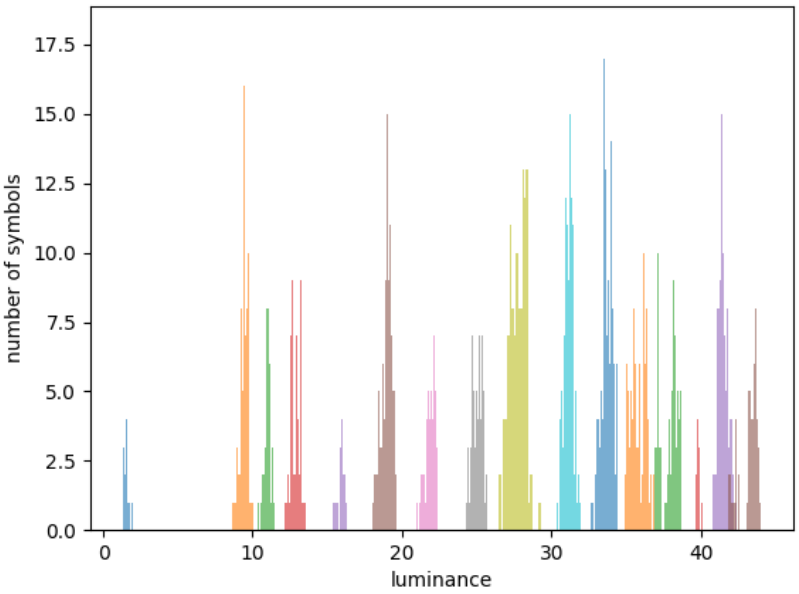


(a)

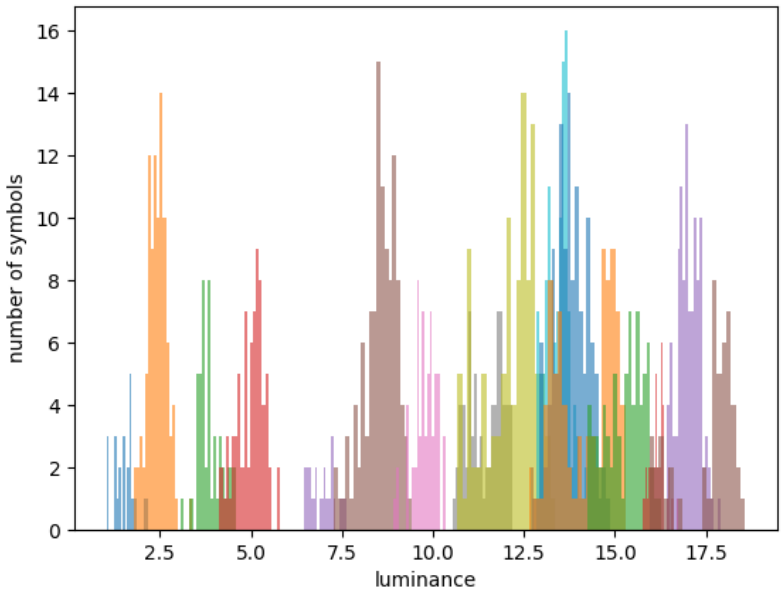


(b)

Figure 4.5: Luminance distributions for PWM-16 at a distance of 20 m. The aperture values are (a) f/2. (b) f/4. (c) f/8. (d) f/16.



(c)



(d)

Figure 4.5: (*continued*) Luminance distributions for PWM-16 at a distance of 20 m. The aperture values are (a) f/2. (b) f/4. (c) f/8. (d) f/16.

4.3.3 Performance Evaluation for PWM-8

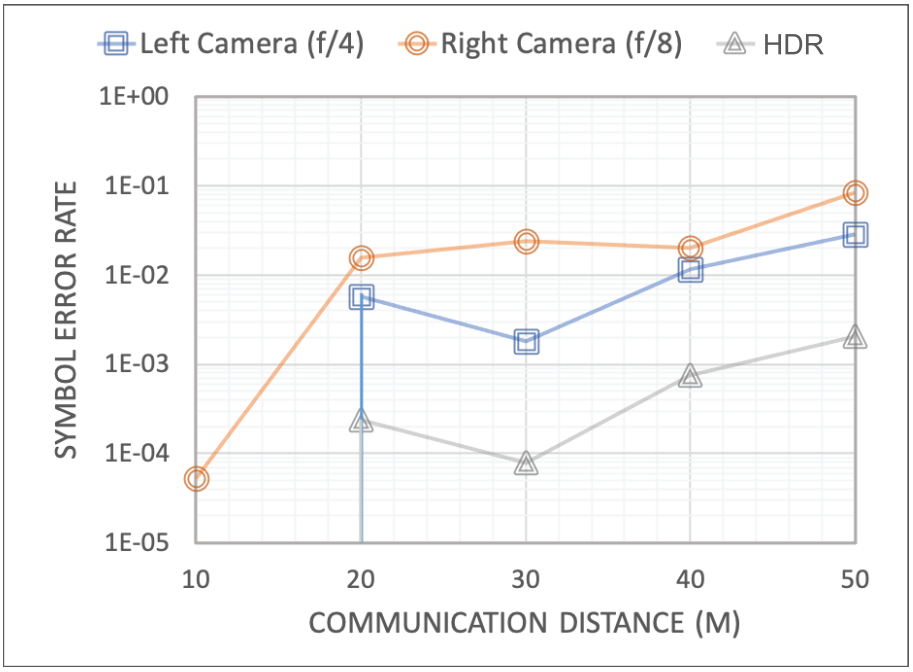
Figures 4.6 illustrate the SER performance under PWM-8, while Figure 4.7 provides a throughput result. The SER curves in the figures represent different components: rectangles represent the SER of the left camera, circles represent the SER of the right camera, and triangles represent the SER of the combining method. SER is calculated by dividing the number of error symbols by the number of transmitted symbols.

The HDR combining is compared to diversity combining. The HDR method involved using different aperture parameters for the left and right cameras. In contrast, the diversity combining approach refers to the demodulation method used in Chapter 3, in which both cameras have the same aperture settings.

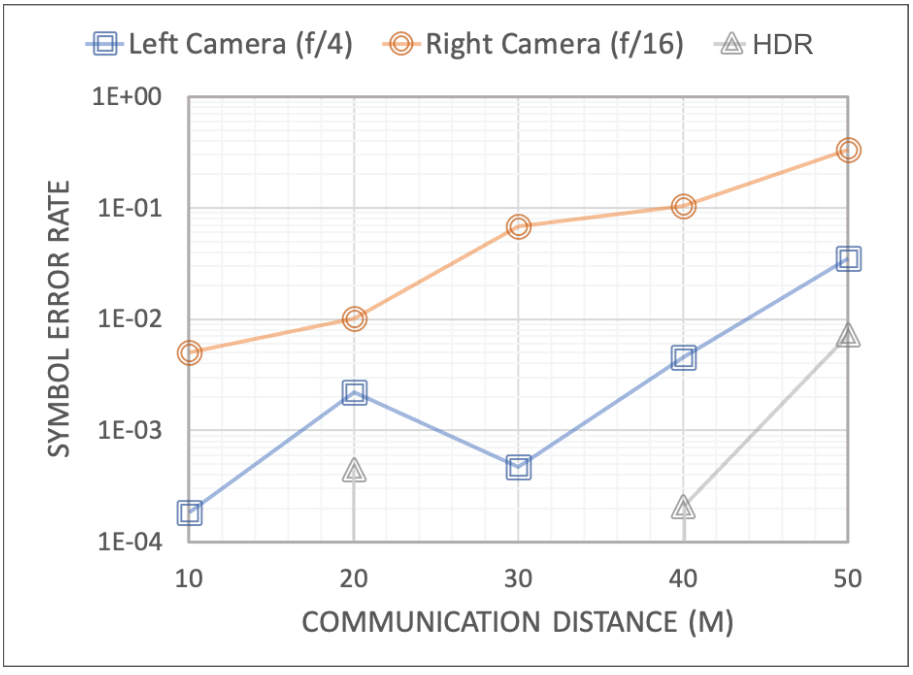
For Figure 4.6 (a), it is evident that both the left camera and HDR achieved error-free transmission at a distance of 10 m. Moving on to Figure 4.6 (b), the HDR method demonstrated error-free transmission at 10 m and 30 m distances. Figure 4.6 (c) reveals that error-free transmission was achieved by the left camera, the right camera, and HDR at a distance of 10 m.

The SER results indicate that the HDR method outperforms using a single camera. When comparing the experimental results at different apertures for the right camera, it becomes apparent that increasing the aperture difference between the two cameras results in a greater disparity in SER. When the cameras have the same aperture, the received LEDs by both cameras exhibit similarities. However, as the camera's aperture difference increases, the LEDs' luminance saturation and diffusion levels on the image plane differ, leading to variations in SER.

Furthermore, the throughput was measured with three different aperture settings, as

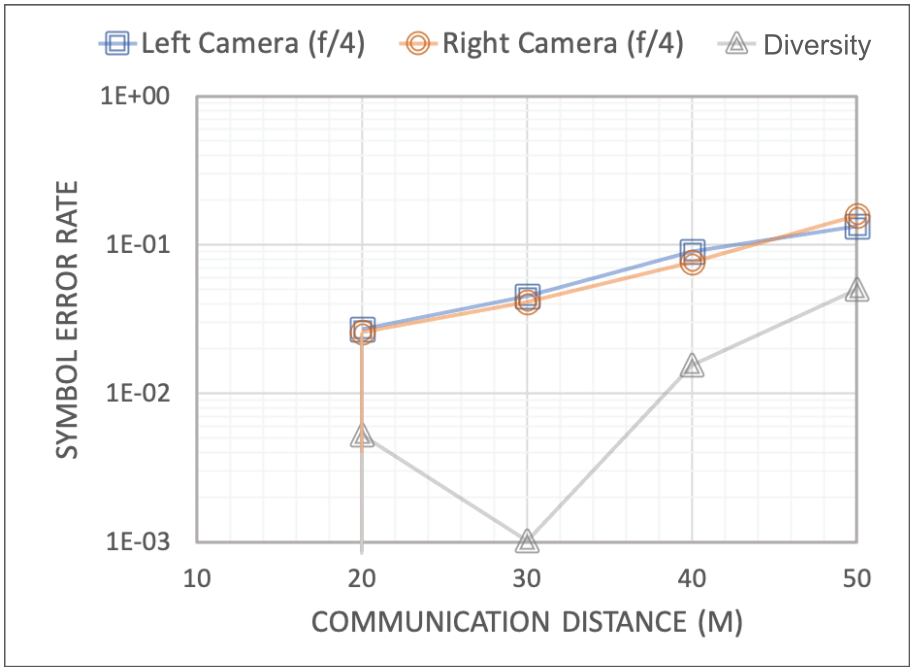


(a)



(b)

Figure 4.6: Symbol error rate result for PWM-8. The left camera is f/4 for all the cases. The right camera is (a) f/8. (b)f/16. (c) f/4.



(c)

Figure 4.6: (*continued*)Symbol error rate for PWM-8. The left camera is f/4 for all the cases. The right camera is (a) f/8. (b)f/16. (c) f/4.

illustrated in Figure 4.7. The triangles correspond to f/4 (left camera) and f/8 (right camera), rectangles represent f/4 (left camera) and f/16 (right camera), and circles indicate f/4 (left camera) and f/4 (right camera).

In this study, throughput is measured in kilobits per second (kbps), defined by the successfully delivered bit per second, given by

$$Throughput = \log_2(M) \times \frac{N \times F}{\Delta t} \times (1 - E) \times \frac{1}{1000} \tag{4.4}$$

where F is the number of data frames, N is the number of LEDs, E is the SER, and Δt is the transmission time. Please note that (4.4) is not a generally correct definition for

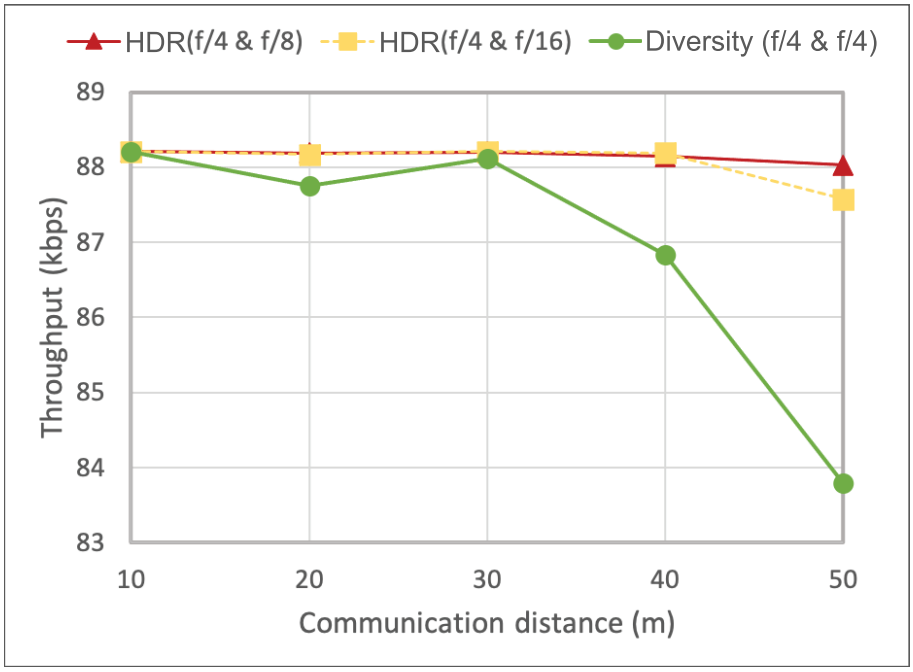


Figure 4.7: Throughput measurement result for PWM-8.

throughput. The correct formula could involve using the packet error rate instead of SER and using error detection codes to detect errors, or it could be based on the Shannon channel capacity to calculate throughput. However, a simpler calculation method was used for this dissertation, where throughput is only used as a numerical value for comparing HDR and diversity reception.

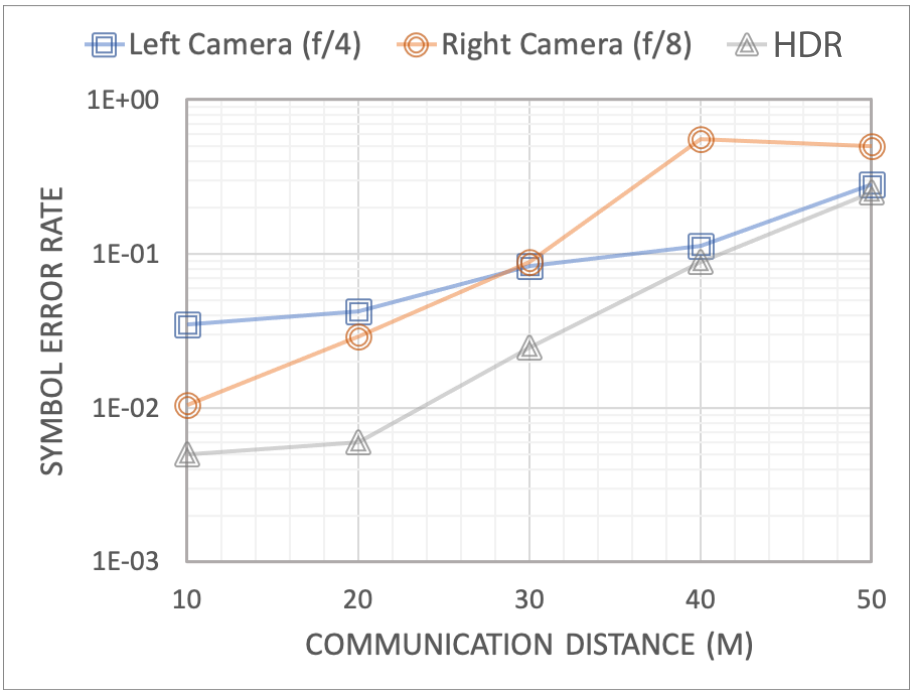
A throughput of 87.57 kbps was achieved at 50 m using HDR (f/4 & f/8). These results indicate that the HDR method, which involves varying camera apertures, enables transmitting a higher number of symbols per second than the diversity method. Interestingly, the throughput remains consistent for the cases where the right camera has apertures of f/8 and f/16. This similarity arises because the HDR image obtained in both cases contains

only two exposures, leading to an equivalent quantity of additional luminance information.

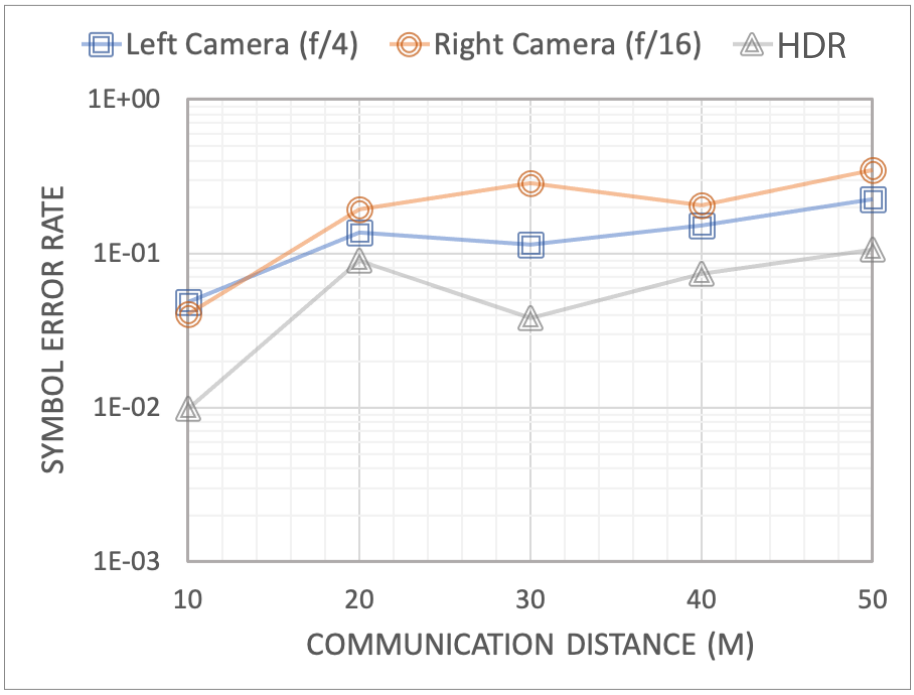
4.3.4 Performance Evaluation for PWM-16

Figure 4.8 illustrate the SER performances of PWM-16 for various aperture values. Additionally, Figure 4.9 presents the throughput performance. Similar to PWM-8, the employment of the HDR method yields superior SER performance compared to using a single camera. However, the overall SER of PWM-16 is inferior to PWM-8 due to the increased number of luminance levels. Despite both PWM-8 and PWM-16 having the same image bit depth, the expanded number of luminance levels in PWM-16 results in narrower differences in pixel values between adjacent levels. Regarding throughput, the HDR approach does not significantly enhance performance, indicating that the impact of different exposures is less significant compared to increased luminance levels when determining symbols. Within a 40 m range, PWM-16 achieves higher throughput than PWM-8. The lowest throughput of PWM-16 at 40 m is recorded as 107.26 kbps, while PWM-8 reaches a maximum throughput of 88.19 kbps. However, it is important to note that the SER of PWM-16 at 40 m is one hundred times higher than that of PWM-8, signifying a significant degradation in signal quality. When considering the trade-off between SER and throughput, the selection between PWM-8 and PWM-16 should be carefully evaluated.

Table 4.2 compares the throughput with different reception methods and PWM techniques. PWM-16 achieves the highest throughput with f/4 and f/16, while PWM-8 performs second best with f/4 and f/8.

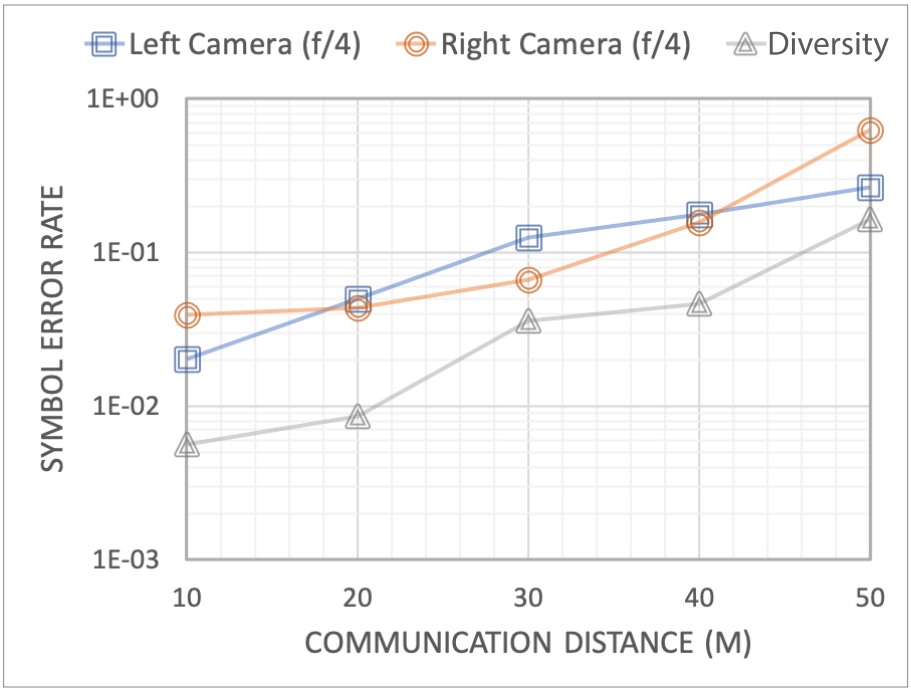


(a)



(b)

Figure 4.8: Symbol error rate for PWM-16. The left camera is f/4 for all the cases. The right camera is (a) f/8. (b)f/16. (c) f/4.



(c)

Figure 4.8: (*continued*)Symbol error rate for PWM-16. The left camera is f/4 for all the cases. The right camera is (a) f/8. (b)f/16. (c) f/4.

4.4 Summary of Chapter 4

This chapter explores the utilization of PWM with multiple luminance levels to enhance the data rate of ISC-ITS. However, the presence of pixel saturation poses challenges for PWM demodulation. The objective of this study is to enable high-level PWM data transmission. To achieve this, HDR images are employed to extend the dynamic range of LED luminance, and the HDR combining technique is applied for PWM demodulation. The results of the field trial experiment demonstrate that a throughput exceeding 87.5 kbps

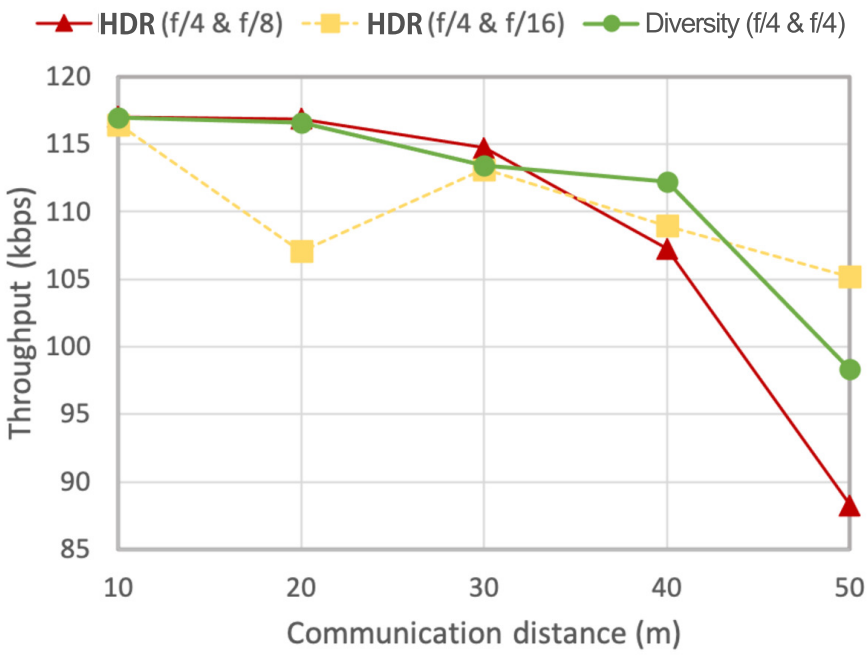


Figure 4.9: Throughput measurement result for PWM-16.

can be achieved within a communication distance of 50 m.

Table 4.2: Compare the minimal throughput of each reception method.

Left Camera	**Right Camera**	**Reception Method**	**PWM[1]-8**	**PWM-16**
N/A	f/4	Single Reception	74.29 kbps	44.4kbps
	f/8		80.71 kbps	58.88 kbps
	f/16		58.78 kbps	76.64 kbps
f/4	N/A	Single Reception	80.71 kbps	58.88 kbps
	f/4	Diversity Reception	83.79 kbps	98.33 kbps
	f/8	HDR Reception[2]	88.02 kbps	88.25 kbps
	f/16	HDR Reception	87.57 kbps	105.17 kbps

[1] PWM: pulse-width modulation

[2] HDR: high dynamic range

Chapter 5

Conclusion

5.1 Summary of this dissertation

This study investigates simultaneous communication and ranging systems for ISC. The system should meet specific requirements, including low latency, robust transmission, and high data rates to ensure efficient and effective communication and ranging performance. In addition, accurate ranging estimation is crucial for obtaining position information and facilitating vehicle navigation. The main issue in this research is the limited dynamic range of the camera, i.e., the LED saturation problem. This study aspires to solve the LED saturation problem through the bicubic interpolation image processing algorithm and HDR implementation using two cameras.

Chapter 1 gives an overview of VLC and introduces the background of this study. Its in-depth background survey is useful for subsequent research, including the survey of the transmitter and receiver, and the survey of the communication and ranging methods.

Chapter 2 provides an in-depth explanation of the system model used in this study. The

stereo camera approach to realize simultaneous communication and ranging is explained in detail.

Chapter 3 describes how bicubic interpolation can be used to cope with the problem of LED saturation in simultaneous VLC and ranging systems. This study found that the LED brightness level significantly affects the disparity estimation in the VLC stereo ranging system. With the use of bicubic interpolation, the error caused by saturation is minimized, which improves the ranging performance and obtains an estimation range with sub-pixel accuracy. The experimental results show that the ranging error of the proposed method is less than 0.5 m over a communication distance of 60 m, which is better than the traditional POC method without bicubic interpolation.

Chapter 4 presents the concept of HDR, a method of utilizing two high-speed cameras with different exposure settings as receivers for an ISC system. Multiple brightness levels of PWM can increase the data transfer rate of ISC-ITS. However, when the PWM level increases, it becomes more challenging to perform PWM demodulation. However, when the PWM level increases, it becomes more challenging to perform PWM demodulation. This study uses high-level PWM to transmit data. HDR images are used to extend the dynamic range of LED brightness. PWM demodulation used HDR combinations. Experimental results show that a throughput higher than 87.5 kbps can be achieved over a communication distance of 50 m.

5.2 Future Developments

This study proposed an HDR combining technique to compensate for the limited dynamic range of LED luminance in ISC systems. While this approach has achieved a high data

rate, there are important areas for future research and development. On the one hand, the current HDR combining technique does not incorporate simultaneous ranging functionality. To address this, exploring the ranging function using HDR images and considering it a topic for future discussion is crucial. On the other hand, in future work, the utilization of luminance information from different exposures to improve the accuracy of symbol determination needs to be considered. This may involve adjusting the weights assigned to each pixel during LED luminance extraction.

By applying this simultaneous VLC and ranging system, high-speed information and high-precision positioning can be obtained because the effect of the low dynamic range of the camera is avoided. 3D depth information obtained through ranging is integrated with road traffic information and map data that can be transmitted through VLC from LED light sources on the road. This integration generates a detailed 3-D map of the vehicle's surroundings, thus contributing to safety and autonomous driver assistance. In addition, the application of VLC in smart city infrastructures can transform urban environments. Smart street lighting systems equipped with VLC technology provide efficient lighting and act as data communication nodes to transmit real-time traffic information, environmental data, and emergency alerts. Combining distance estimation with road traffic information and map data can further contribute to developing advanced traffic management systems that improve the safety and efficiency of urban transportation.

From a broader future development perspective, integrating VLC with ranging will have great transformative potential for a wide range of industries. In addition to applications in ITS, the technology can also be used in areas such as robotics and remote sensing for environmental monitoring. In the field of robotics, high-speed data transmission and high-precision distance estimation achieved by the proposed system can significantly en-

hance the sensing capability of the robot platform. Robots equipped with VLC-based technology can accurately navigate through complex environments, safely interact with humans, and precisely perform delicate tasks. In the field of remote sensing and environmental monitoring, VLC using drones can help accurately monitor land, water bodies, and vegetation, supporting applications in agriculture, environmental protection, and disaster management.

In conclusion, the proposed system in VLC and ranging holds immense promise in revolutionizing various industries. By harnessing the power of this technology, we can envision a future with smarter, safer, and more interactive environments in transportation, robotics, smart city infrastructure, etc. The continuous exploration and advancement of VLC will undoubtedly pave the way for transformative innovations and enrich our daily lives.

Reference

[1] G. Pang, T. Kwan, C.-H. Chan, and H. Liu, "LED traffic light as a communications device," in *Proceedings 199 IEEE/IEEJ/JSAI International Conference on Intelligent Transportation Systems (Cat. No.99TH8383)*, pp. 788–793, 1999.

[2] T. Komine and M. Nakagawa, "Fundamental analysis for visible-light communication system using LED lights," *IEEE Transactions on Consumer Electronics*, vol. 50, pp. 100–107, Feb 2004.

[3] R. Huang and T. Yamazato, "A review on image sensor communication and its applications to vehicles," *Photonics*, vol. 10, no. 6, 2023.

[4] J. Ohta, *Smart CMOS Image Sensors and Applications*. CRC Press, 5 2020.

[5] G. K. H. Pang and H. H. S. Liu, "LED location beacon system based on processing of digital images," *IEEE Transactions on Intelligent Transportation Systems*, vol. 2, pp. 135–150, Sep 2001.

[6] C. Danakis, M. Afgani, G. Povey, I. Underwood, and H. Haas, "Using a CMOS camera sensor for visible light communication," in *2012 IEEE Globecom Workshops*, pp. 1244–1248, 2012.

[7] T. Yamazato, I. Takai, H. Okada, T. Fujii, T. Yendo, S. Arai, M. Andoh, T. Harada, K. Yasutomi, K. Kagawa, and S. Kawahito, "Image-sensor-based visible light communication for automotive applications," *IEEE Communications Magazine*, vol. 52, pp. 88–97, July 2014.

[8] T. Yamazato, M. Kinoshita, S. Arai, E. Souke, T. Yendo, T. Fujii, K. Kamakura, and H. Okada, "Vehicle motion and pixel illumination modeling for image sensor based visible light communication," *IEEE Journal on Selected Areas in Communications*, vol. 33, pp. 1793–1805, Sept 2015.

[9] M. Kinoshita, T. Yamazato, H. Okada, T. Fujii, S. Arai, T. Yendo, and K. Kamakura, "Motion modeling of mobile transmitter for image sensor based I2V-VLC, V2I-VLC, and V2V-VLC," in *Globecom Workshops (GC Wkshps), 2014*, pp. 450–455, Dec 2014.

[10] A. D. Griffiths, J. Herrnsdorf, M. J. Strain, and M. D. Dawson, "Scalable visible light communications with a micro-led array projector and high-speed smartphone camera," *Opt. Express*, vol. 27, pp. 15585–15594, May 2019.

[11] P. Chavez‐Burbano, S. Vitek, S. Teli, V. Guerra, J. Rabadan, R. Perez‐Jimenez, and S. Zvanovec, "Optical camera communication system for internet of things based on organic light emitting diodes," *Electronics Letters*, vol. 55, pp. 334–336, 3 2019.

[12] T. Arisue, T. Yamazato, H. Okada, T. Fujii, M. Kinoshita, K. Kamakura, S. Arai, and T. Yendo, "BER measurement for transmission pattern design of ITS image sensor communication using DMD projector," in *2020 IEEE 17th Annual Consumer Communications & Networking Conference (CCNC)*, pp. 1–6, 2020.

[13] S. Zhong, Y. Zhu, X. Chi, H. Shi, H. Sun, and S. Wang, "Optical lensless-camera communications aided by neural network," *Applied Sciences*, vol. 9, p. 3238, 8 2019.

[14] F. Ahmed, M. H. Conde, P. L. Martinez, T. Kerstein, and B. Buxbaum, "Pseudo-passive time-of-flight imaging: Simultaneous illumination, communication, and 3d sensing," *IEEE Sensors Journal*, vol. 22, pp. 21218–21231, 11 2022.

[15] P. Pathak, X. Feng, P. Hu, and P. Mohapatra, "Visible light communication, networking, and sensing: A survey, potential and challenges," *IEEE Communications Surveys and Tutorials*, vol. 17, pp. 2047–2077, 2015.

[16] A. Jovicic, J. Li, and T. Richardson, "Visible light communication: Opportunities, challenges and the path to market," *IEEE Communications Magazine*, vol. 51, pp. 26–32, 2013.

[17] D. Karunatilaka, F. Zafar, V. Kalavally, and R. Parthiban, "LED based indoor visible light communications: State of the art," *IEEE Communications Surveys and Tutorials*, vol. 17, pp. 1649–1678, 2015.

[18] H. Haas, L. Yin, Y. Wang, and C. Chen, "What is LiFi?," *Journal of Lightwave Technology*, vol. 34, pp. 1533–1544, March 2016.

[19] A. G. Bell, "On the production and reproduction of sound by light," *American Journal of Science*, vol. s3-20, pp. 305–324, 10 1880.

[20] D. Vanderwater, I.-H. Tan, G. Hofler, D. Defevere, and F. Kish, "High-brightness algainp light emitting diodes," *Proceedings of the IEEE*, vol. 85, no. 11, pp. 1752–1764, 1997.

[21] S. Nakamura, T. M. T. Mukai, and M. S. M. Senoh, "High-power GaN P-N junction blue-light-emitting diodes," *Japanese Journal of Applied Physics*, vol. 30, p. L1998, dec 1991.

[22] Y. Tanaka, S. Haruyama, and M. Nakagawa, "Wireless optical transmissions with white colored LED for wireless home links," in *11th IEEE International Symposium on Personal Indoor and Mobile Radio Communications. PIMRC 2000. Proceedings (Cat. No.00TH8525)*, vol. 2, pp. 1325–1329 vol.2, 2000.

[23] S. Haruyama and T. Yamazato, "Image sensor based visible light communication," in *Visible Light Communication*, pp. 181–205, Cambridge University Press, 2 2015.

[24] K. Kamakura, "Image sensors meet LEDs," *IEICE Transactions on Communications*, vol. E100.B, pp. 917–925, 2017.

[25] T. Nguyen, A. Islam, T. Yamazato, and Y. M. Jang, "Technical issues on IEEE 802.15.7m image sensor communication standardization," *IEEE Communications Magazine*, vol. 56, no. 2, pp. 213–218, 2018.

[26] T. Nguyen, A. Islam, T. Hossan, and Y. M. Jang, "Current status and performance analysis of optical camera communication technologies for 5G networks," *IEEE Access*, vol. 5, pp. 4574–4594, 2017.

[27] M. K. Hasan, M. O. Ali, M. H. Rahman, M. Z. Chowdhury, and Y. M. Jang, "Optical camera communication in vehicular applications: A review," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 6260–6281, 2022.

[28] "IEEE standard for local and metropolitan area networks–part 15.7: Short-range wireless optical communication using visible light," *IEEE Std 802.15.7-2011*, pp. 1–309, 2011.

[29] IEEE Std 802.15.7-2018 (Revision of IEEE Std 802.15.7-2011), "IEEE standard for local and metropolitan area networks–part 15.7: Short-range optical wireless communications," 2019.

[30] "ISO 22738:2020 Intelligent transport systems — Localized communications — Optical camera communication." "https://www.iso.org/standard/73769.html", 2020. [Online;].

[31] S.-H. Yang, H.-S. Kim, Y.-H. Son, and S.-K. Han, "Three-dimensional visible light indoor localization using AOA and RSS with multiple optical receivers," *Journal of Lightwave Technology*, vol. 32, no. 14, pp. 2480–2485, 2014.

[32] R. Bian, I. Tavakkolnia, and H. Haas, "15.73 Gb/s visible light communication with off-the-shelf LEDs," *Journal of Lightwave Technology*, vol. 37, no. 10, pp. 2418–2424, 2019.

[33] D. Tsonev, S. Videv, and H. Haas, "Towards a 100 Gb/s visible light wireless access network," *Opt. Express*, vol. 23, pp. 1627–1637, Jan 2015.

[34] T. Yamazato, A. Ohmura, H. Okada, T. Fujii, T. Yendo, S. Arai, and K. Kamakura, "Range estimation scheme for integrated i2v-vlc using a high-speed image sensor," in *2016 IEEE International Conference on Communications Workshops (ICC)*, May 2017.

[35] K. Abboud, H. A. Omar, and W. Zhuang, "Interworking of DSRC and cellular network technologies for V2X communications: A survey," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 12, pp. 9457–9470, 2016.

[36] A. Lei, H. Cruickshank, Y. Cao, P. Asuquo, C. P. A. Ogah, and Z. Sun, "Blockchain-based dynamic key management for heterogeneous intelligent transportation systems," *IEEE Internet of Things Journal*, vol. 4, no. 6, pp. 1832–1843, 2017.

[37] E. Yurtsever, J. Lambert, A. Carballo, and K. Takeda, "A survey of autonomous driving: Common practices and emerging technologies," *IEEE Access*, vol. 8, pp. 58443–58469, 2020.

[38] S. Kamegawa, M. Kinoshita, T. Yamazato, H. Okada, T. Fujii, K. Kamakura, T. Yendo, and S. Arai, "Performance evaluation of precoded pulse width modulation for image sensor communication," in *2018 IEEE Globecom Workshops (GC Wkshps)*, IEEE, dec 2018.

[39] V. T. B. Tram and M. Yoo, "Vehicle-to-vehicle distance estimation using a low-resolution camera based on visible light communications," *IEEE Access*, vol. 6, pp. 4521–4527, Jan. 2018.

[40] T. Nagura, T. Yamazato, M. Katayama, T. Yendo, T. Fujii, and H. Okada, "Tracking an LED array transmitter for visible light communications in the driving situation," in *2010 7th International Symposium on Wireless Communication Systems (ISWCS)*, pp. 765–769, Sept 2010.

[41] Apple, "iphone 14 specifications," 2023. https://www.apple.com/iphone-14/specs/.

[42] J. B. Kenney, "Dedicated short-range communications (DSRC) standards in the united states," *Proceedings of the IEEE*, vol. 99, no. 7, pp. 1162–1182, 2011.

[43] ETSI TR 102 863 V1.1.1 (2011-06), "Intelligent transport systems (ITS); vehicular communications; basic set of applications; local dynamic map (LDM); rationale for and guidance on standardization." Standard, June 2011.

[44] H. Wymeersch, J. Lien, and M. Z. Win, "Cooperative localization in wireless networks," *Proceedings of the IEEE*, vol. 97, pp. 427–450, 2 2009.

[45] K. Takahashi, K. Kamakura, M. Kinoshita, and T. Yamazato, "Nonlinear transform for parallel transmission for image-sensor-based visible light communication," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, pp. 1–5, 2020.

[46] S. Kibe, K. Kamakura, and T. Yamazato, "Parallel transmission using M-Point DFT for image-sensor-based visible light communication," in *2018 IEEE Globecom Workshops (GC Wkshps)*, pp. 1–6, 2018.

[47] S. Nishimoto, T. Nagura, T. Yamazato, T. Yendo, T. Fujii, H. Okada, and S. Arai, "Overlay coding for road-to-vehicle visible light communication using LED array and high-speed camera," in *14th International IEEE Conference on Intelligent Transportation Systems (ITSC 2011)*, pp. 1704–1709, Oct 2011.

[48] S. Nishimoto, T. Yamazato, H. Okada, T. Fujii, T. Yendo, and S. Arai, "High-speed transmission of overlay coding for road-to-vehicle visible light communication using LED array and high-speed camera," in *IEEE Globecom Workshops (OWC 2012)*, pp. 1234–1238, Dec 2012.

[49] T.-H. Do and M. Yoo, "Multiple exposure coding for short and long dual transmission in vehicle optical camera communication," *IEEE Access*, vol. 7, pp. 35148–35161, 2019.

[50] R. Roberts, "A MIMO protocol for camera communications (CamCom) using undersampLED frequency shift ON-OFF keying (UFSOOK)," pp. 1052–1057, 2013.

[51] H. Aoyama and M. Oshima, "Line scan sampling for visible light communication: Theory and practice," in *2015 IEEE International Conference on Communications (ICC)*, pp. 5060–5065, 2015.

[52] M. Kinoshita, K. Kamakura, T. Yamazato, H. Okada, T. Fujii, S. Arai, and T. Yendo, "Stereo ranging method using LED transmitter for visible light communication," pp. 1–6, IEEE, 12 2019.

[53] M. Kinoshita, T. Yamazato, H. Okada, T. Fujii, S. Arai, T. Yendo, and K. Kamakura, "Experimental evaluation of visible light communication performance applying pixel selection/maximal ratio combining reception by high-speed stereo camera," *IEICE Transactions on Communications (Japanese Edition)*, vol. J102-B, pp. 90–97, Feb. 2019.

[54] A. Kaehler and G. Bradski, *Learning OpenCV 3: Computer Vision in C++ with the OpenCV Library*. O'Reilly Media, Inc., 1st ed., 2016.

[55] R. Huang, M. Kinoshita, T. Yamazato, H. Okada, K. Kamakura, S. Arai, T. Yendo, and T. Fujii, "Performance evaluation of range estimation for image sensor communication using phase-only correlation," in *2020 IEEE Globecom Workshops (GC Wkshps*, pp. 1–6, 2020.

[56] M. S. Ifthekhar, N. Saha, and Y. M. Jang, "Stereo-vision-based cooperative-vehicle positioning using OCC and neural networks," *Optics Communications*, vol. 352, pp. 166–180, 2015.

[57] T. Do and M. Yoo, "Visible light communication based vehicle positioning using LED street light and rolling shutter cmos sensors," *Optics Communications*, vol. 407, pp. 112–126, 2018.

[58] B. W. Kim and S.-Y. Jung, "Vehicle positioning scheme using V2V and V2I visible light communications," in *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, pp. 1–5, 2016.

[59] M. T. Hossan, M. Z. Chowdhury, M. K. Hasan, M. Shahjalal, T. Nguyen, N. T. Le, and Y. M. Jang, "A new vehicle localization scheme based on combined optical camera communication and photogrammetry," *Mobile Information Systems*, vol. 2018, pp. 1–14, 2018.

[60] J. He, K. Tang, J. He, and J. Shi, "Effective vehicle-to-vehicle positioning method using monocular camera based on vlc," *Opt. Express*, vol. 28, pp. 4433–4443, Feb 2020.

[61] T. Yamazato, A. Ohmura, H. Okada, T. Fujii, T. Yendo, S. Arai, and K. Kamakura, "Range estimation scheme for integrated I2V-VLC using a high-speed image sensor," in *2016 IEEE International Conference on Communications Workshops (ICC)*, pp. 326–330, May 2016.

[62] J. He and B. Zhou, "A deep learning-assisted visible light positioning scheme for vehicles with image sensor," *IEEE Photonics Journal*, vol. 14, pp. 1–7, 8 2022.

[63] A. Islam, M. T. Hossan, and Y. M. Jang, "Convolutional neural networkscheme–based optical camera communication system for intelligent internet of vehicles," *International Journal of Distributed Sensor Networks*, vol. 14, no. 4, p. 1550147718770153, 2018.

[64] Z. Li, A. Liao, L.and Wang, and et al., "Vehicular optical ranging and communication system," *J Wireless Com Network*, vol. 190, no. 3, pp. 1939–1962, 2015.

[65] B. Béchadergue, L. Chassagne, and H. Guan, "Simultaneous visible light communication and distance measurement based on the automotive lighting," *IEEE Trans. Intell. Veh.*, vol. 4, no. 4, pp. 532–547, 2019.

[66] I. Takai, T. Harada, M. Andoh, K. Yasutomi, K. Kagawa, and S. Kawahito, "Optical vehicle-to-vehicle communication system using LED transmitter and camera receiver," *IEEE Photonics Journal*, vol. 6, no. 5, pp. 1–14, 2014.

[67] M. Kinoshita, T. Yamazato, H. Okada, T. Fujii, S. Arai, T. Yendo, and K. Kamakura, "A comparison of reception methods for visible light communication using high-speed stereo cameras," in *2018 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–5, 2018.

[68] I. Takai, T. Harada, M. Andoh, K. Yasutomi, K. Kagawa, and S. Kawahito, "Optical vehicle-to-vehicle communication system using LED transmitter and camera receiver," *IEEE Photonics Journal*, vol. 6, pp. 1–14, Oct 2014.

[69] R. Huang, M. Kinoshita, T. Yamazato, H. Okada, K. Kamakura, S. Arai, T. Yendo, and T. Fujii, "Simultaneous visible light communication and ranging using high-speed stereo cameras based on bicubic interpolation considering multi-level pulse-

width modulation," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E106.A, no. 7, pp. 990–997, 2022.

[70] N. B. Hassan, Z. Ghassemlooy, S. Zvánovec, M. Biagi, A. M. Vegni, M. Zhang, and Y. Huang, "Interference cancellation in mimo nlos optical-camera-communication-based intelligent transport systems.," *Applied optics*, vol. 58 34, pp. 9384–9391, 2019.

[71] S. Usui, T. Yamazato, S. Arai, T. Yendo, T. Fujii, and H. Okada, "Utilization of spatio-temporal image for LED array acquisition in road to vehicle visible light communication," in *20th World Congress on Intelligent Transport Systems*, Oct. 2013.

[72] R. Huang, M. Kinoshita, T. Yamazato, H. Okada, K. Kamakura, S. Arai, T. Yendo, and T. Fujii, "Performance evaluation of range estimation for image sensor communication using phase-only correlation," *2020 IEEE Global Communications Conference (GLOBECOM) Workshops*, Dec. 2020.

[73] K. Saneyoshi, "Drive assist system using stereo image recognition," in *Proceedings of Conference on Intelligent Vehicles*, pp. 230–235, 1996.

[74] D. Nister, O. Naroditsky, and J. Bergen, "Visual odometry," in *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, vol. 1, pp. I–I, 2004.

[75] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 911–916, 2000.

[76] R. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 29, no. 6, pp. 1153–1160, 1981.

[77] C. D. Kuglin and D. C. Hines, "The phase correlation image alignment method," in *Proc. Int. Conf. Cybernetics Society*, p. 163–165, 1975.

[78] H. Foroosh, J. B. Zerubia, and M. Berthod, "Extension of phase correlation to subpixel registration," *IEEE Transactions on Image Processing*, vol. 11, pp. 911–916, Mar. 2002.

[79] K. Takita, T. Aoki, Y. Sasaki, T. Higuchi, and K. Kobayashi, "High-accuracy subpixel image registration based on phase-only correlation," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E86-A, 08 2003.

[80] S. Mann and R. W. Picard, "On being 'undigital' with digital cameras: Extending dynamic range by combining differently exposed pictures," in *PROCEEDINGS OF IS&T*, pp. 442–448, 1995.

[81] S. Suzuki and K. be, "Topological structural analysis of digitized binary images by border following," *Computer Vision, Graphics, and Image Processing*, vol. 30, no. 1, pp. 32–46, 1985.

List of Symbols

Symbol	Parameter
P	the 3-D position of the LED at the world coordinates
X	horizontal distance of the LED from the left camera position at the world coordinates
Y	vertical distance of the LED from the left camera position at the world coordinates
Z	distance (or depth) of LED from the two cameras at the world coordinates
b	distance between the left and right cameras
P_{left}	the two-dimensional position of the LED at the image coordinates of the left camera
x_l	the horizontal position of the LED at the image coordinates of the left camera
v	the vertical position of the LED at the image coordinates of the left camera

Symbol	Parameter
P_{right}	the two-dimensional position of the LED at the image coordinates of right camera
x_r	the horizontal position of the LED at the image coordinates of the right camera
f	the focal length of the cameras
ρ	the parameter to convert disparity to the actual size of a single pixel
d	disparity
M	the maximal number of PWM levels
V_{LED}	LED forward voltage
T	the time for one period
D	pulse duration time, where $D \in \{0, T/(M-1), 2T/(M-1), \ldots, T\}, (M > 2)$
x	the horizontal position of a point on the received image
y	the vertical position of a point on the received image
Δ_x	the decimal subpixel value on the horizontal axis of the received image corresponds to the interpolation image
Δ_y	the decimal subpixel value on the vertical axis of the received image corresponds to the interpolation image
x_{bic}	the horizontal position of corresponding point of (x, y) on the interpolation image
y_{bic}	the vertical position of corresponding point of (x, y) on the interpolation image
α	magnification factor

Symbol	Parameter
G	the received image
G_{bic}	the interpolation image
v	the position difference between the interpolation point and the current pixel
h	the position difference between the interpolation point and the current pixel
u	the position difference between the interpolation point and the current pixel
S	the position difference between the interpolation point and the current pixel
G_{left}	interpolated signals of the images captured by the left camera
G_{right}	interpolated signals of the images captured by the right camera
$\Delta(x_{bic})$	the decimal subpixel value on the horizontal axis of the received image corresponds to the interpolation image
$\Delta(y_{bic})$	the decimal subpixel value on the vertical axis of the received image corresponds to the interpolation image
$C(x_{bic}, y_{bic})$	the cross-power spectrum of $G_{left}(x_{bic}, y_{bic})$
$c(x_{bic}, y_{bic})$	inverse discrete Fourier transform of the cross-power spectrum of $G_{left}(x_{bic}, y_{bic})$
δ_K	a Kronecker's delta function
W	the width of the image
H	the height of the image
δ_1	the subpixel image displacement in the x-axis
δ_2	the subpixel image displacement in the y-axis

Symbol	Parameter
$\gamma(x_{bic})$	the image displacement in the subpixel level on the x-axis
$w_l^{s'}$	channel factors of the left camera
$w_r^{s'}$	channel factors of the right camera
s'	the transmitted symbol in the training sequence
$N_{s'}$	the data length of training sequence for symbol s'
$L_l^{s'}[n]$	the n-th received luminance values of symbol s' in training sequence for the left camera
$L_r^{s'}[n]$	the n-th received luminance values of symbol s' in training sequence for the right camera
$\hat{s}$	estimated symbol
L_l	the received luminance values in the data sequence of the left camera
L_r	the received luminance values in the data sequence of the right camera
s	a transmission symbol in the data sequence
D_l	the received luminance values in the data sequence of the left camera
D_r	the received luminance values in the data sequence of the right camera
h_l^s	channel factors of the left camera
h_r^s	channel factors of the right camera
N	number of LEDs
F	number of data frames
E	symbol error rate

Symbol	Parameter
Δt	the transmission time

Acknowledgements

When I came to Japan five years ago, I never imagined that I would end up studying for a doctoral degree in Japan. Pursuing my master's and doctoral degrees overseas has been an invaluable experience that has allowed me to learn about different cultures. I am deeply grateful to the many people and organizations that have supported me along the way, without whom it would not have been possible to complete my research.

First of all, I would like to express my heartfelt thanks and gratitude to my supervisor Dr. Takaya Yamazato, Professor at the Institute of Liberal Arts and Sciences Education Planning & Development Center, Nagoya University, for his patience and guidance over the past five years. Thank you for your guidance and encouragement, which has greatly benefited both my academic and personal growth. Your support has been a driving force for me, and I will always cherish it.

Second, I would like to thank my dissertation committee. Dr. Massaki Katayama, Professor at the Institute of Materials and Systems for Sustainability, Nagoya University, Dr. Hiraku Okada, Associate Professor at the Institute of Materials and Systems for Sustainability, Nagoya University, Dr. Fujii Toshiaki, Professor at Nagoya University, and Dr. Wataru Chujo, Professor at Meijo University, for reviewing this paper. I would like to thank them for taking out of their time to review this dissertation and for their very useful

comments and suggestions on my research.

Next, I would like to express my sincere gratitude to Dr. Masayuki Kinoshita, Associate Professor at Chiba Institute of Technology, for his guidance, comments, and suggestions on my research. I would like to extend my gratitude to Dr. Shintaro Arai of Okayama University of Science, Dr. Koji Kamakura of Chiba Institute of Technology, and Dr. Tomohiro Yendo of Nagaoka University of Technology for their invaluable input, guidance, and collaborative research on visible light communication. I am also grateful to Dr. Chedlia Ben Naila, Assistant Professor at the Institute of Materials and Systems for Sustainability, Nagoya University, and Associate Professor Kentaro Kobayashi of Meijo University, for their useful suggestions and support.

Furthermore, I would like to express my special thanks to the office secretaries, Aiko Ishikawa and Eriko Shiraishi. They listened patiently and understandingly to the difficulties I encountered in an unfamiliar environment and provided valuable guidance and support. Their kindness helped me overcome the challenges of cultural shock.

In addition, I would like to thank the students in the Yamazato and Katayama labs for their support and advice. I would like to express my gratitude to the members of Kashikoi research group: Syunki Kamiya, Kento Nakamura, Takuya Yamamoto, Satoshi Kamegawa, Masahiro Hori, Zhengqiang Tang, and Kosuke Nakano, and the members of the Katayama-Yamazato Laboratory, Xuejuan Zhu, Jinxing Zheng, Dr. Meiyuan Miao, Dr. Yasuo Nakashima, and Mr. Hideki Omote. I am grateful for their participation in the discussions, valuable insights, and important contributions to this research. In particular, I would like to express my special thanks to my tutor, Tomoya Arisue, and my senpai, Dr. Akira Tsujii, who provided a lot of help in both my daily life and academic studies during my initial days in Japan.

I would also like to take this opportunity to thank the Ministry of Education, Culture, Sports, Science and Technology (MEXT) Scholarship, the Nagoya University Interdisciplinary Frontier Fellowship, and the Graduate School of Engineering, Nagoya University, for providing financial support for research fundings and my living expenses.

Last but not least, I would like to thank my family and friends: Xiaoxun Huang, Yuxi Song, Noha Harag, Tina Wang, Astrid Xie, Ying Zou, and Xinqi Chen, for their unwavering patience and support throughout my study.

Research Publications

（Circled items represent references relevant to this study）

I. Journal Paper

	Title	Authority and Date	Authors
1.	Roadside LED Array Acquisition For Road-to-Vehicle Visible Light Communication Using Spatial-Temporal Gradient Values	IEICE Communications Express, vol. 11, no. 8, pp. 462-467, Aug. 2022.	K. Nakamura, R. Huang, Y. Yamazato, M. Kinoshita, K. Kamakura, S.Arai, T. Yendo, T. Fujii
②.	Simultaneous Visible Light Communication and Ranging Using High-Speed Stereo Cameras Based on Bicubic Interpolation Considering Multi-Level Pulse-Width Modulation	IEICE Transactions on Fundamentals of Electronics Communications and Computer Sciences, vol.E106-A, no.7, pp.990-997, Jul. 2023.	R. Huang, M. Kinoshita, T. Yamazato, H. Okada, K. Kamakura, S. Arai, T.Yendo, T. Fujii

	Title	Authority and Date	Authors
③.	A Review on Image Sensor Communication and its Applications to Vehicles	Photonics, vol. 10, no. 6, pp. 617-647, May 2023.	R. Huang, Y. Yamazato,
④.	High-Speed Image Sensor Communication Using High Dynamic Range Combining	IEICE Communications Express, advance online publication, DOI: https://doi.org/10.1587/comex.2023XBL0075	R. Huang, T. Yamazato, M. Kinoshita, H. Okada, K. Kamakura, S. Arai, T. Yendo, T. Fujii

※ IEICE: The Institute of Electronics, Information and Communication Engineers

II. International Conference

	Title	Authority and Date	Authors
①	Calibration method for an integrated ranging and visible light communication system using stereo cameras	International Conference on Materials and Systems for Sustainability (ICMaSS), Nov. 2019.	R. Huang, M. Kinoshita, T. Yamazato, H. Okada, T. Fujii, S. Arai, T. Yendo, K. Kamakura
②	Performance Evaluation of Range Estimation for Image Sensor Communication Using Phase-only Correlation	2020 IEEE Globecom Workshops (GC Wkshps), pp. 1-6, Dec.2020.	R. Huang, M. Kinoshita, T. Yamazato, H. Okada, K. Kamakura, S. Arai, T. Yendo, T. Fujii

	Title	Authority and Date	Authors
3.	Vehicle Distance Measurement based on Visible Light Communication Using Stereo Cameras	2021 IEEE Intelligent Vehicles Symposium (IV), pp. 853-858, Jul.2021.	R. Huang, T. Yamazato, M. Kinoshita, H. Okada, K. Kamakura, S. Arai, T. Yendo, T. Fujii

※ IEEE: The Institute of Electrical and Electronics Engineers

III. Others

	Title	Authority and Date	Authors
1.	Range Estimation for Traffic Light based on Visible Light Communication using Stereo Cameras	Workshop on Optical Wireless Communication for Smart City (OWC2), pp.49-52, Dec. 2019.	R. Huang, T. Yamazato, H. Okada, M. Kinoshita, K. Kamakura, S. Arai, T. Yendo, T. Fujii
②.	Range Estimation of LED Transmitter Using Phase-Only Correlation for ITS-VLC System via High-speed Stereo Cameras	IEICE General Conference, A-9-21, p.100, Mar. 2020	R. Huang, M. Kinoshita, T. Yamazato, H. Okada, K. Kamakura, S. Arai, T. Yendo, T. Fujii

	Title	Authority and Date	Authors
3.	高速二眼カメラを用いた可視光通信による車両前方測距	電子情報通信学会技術研究報告, CS2021-22, pp.33-37, Jul. 2021.	黄瑞怡, 山里敬也, 岡田啓, 木下雅之, 荒井伸太郎, 鎌倉浩嗣, 圓道知博, 藤井俊彰

※ IEICE: The Institute of Electronics, Information and Communication Engineers

IV. Awards

- 電子情報通信学会東海支部令和 4 年度学生研究奨励賞，2023 年 6 月